

QUARTERLY JOURNAL OF
EXPERIMENTAL PSYCHOLOGY

VOL. VIII

1956

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY

Vol. VIII

1956

Part 1

SEX DIFFERENCES IN TACTUAL-KINAESTHETIC AND VISUAL PERCEPTION OF VERTICALITY

BY

CARL IVAR SANDSTRÖM

From the Psychological Laboratory, University of Stockholm

Previous studies on visual and tactual-kinaesthetic perception with head tilted sidewise have shown that the greater deviations are to the left, i.e. the adjustments of a luminous line or a pivoting rod tend to fall counter-clockwise to the perpendicular. This paper shows significant sex difference in tactual-kinaesthetic perception only. Men and women show similar and small deviations to the right with head tilted to the left but, with head tilted to the right, women show clearly greater deviations to the left. The discussion of the visual perception of verticality (the so-called Aubert phenomenon) has been confusing. Our results in the visual experiments show small deviations, and they support Aubert's statement.

I

INTRODUCTION

The research work on the basic problems in this paper originates in a study by Aubert (1861). He stated that a true-vertical luminous line observed in a dark room with head tilted sidewise appears to be displaced in the laterally *opposite* quadrant. If we let a subject with head tilted adjust such a luminous line from a certain position of deviation to the apparent true upright, Aubert's statement implies that the luminous line will be moved to the *same* quadrant as the tilt of the head.

In two recent independent papers (Sandström, 1952; Wapner & Warner, 1952) it has been shown that the corresponding tactual-kinaesthetic perception of verticality is opposite to this so-called Aubert-phenomenon (or A-phenomenon).

In this paper the stress is laid upon sex differences. Such differences have previously been demonstrated within the opto-kinaesthetic field by Sandström (1953) and in related fields; especially by Witkin, who gave a detailed synopsis of his results in 1954.

II

APPARATUS AND PROCEDURE

When studying the tactual-kinaesthetic perception of verticality the task is to adjust a rod, pivoting at its centre, to the apparent upright by using the hands as tactual-kinaesthetic media in an absolutely dark room. Subjects are asked to move one of the hands or both of them up and down along the rod, lightly held between the thumb and the other fingers. The apparatus employed is previously described and reproduced (Sandström, 1952).

For the experiments on visual perception the same apparatus was used with this modification: Of the 38-cm-long metal rod to be used for the adjustment to the upright,

31 cm was painted with luminescent paint which gave a very distinct light in the dark room. The unpainted part at the bottom of the rod was used as a handle for the subject during the adjustment.

Thus, the same psychophysical method (method of reproduction) has been used throughout all the experiments.

In both types of experiment the subject's distance to the rod was about 50 cm.

Experiments 1 and 2 dealt with tactual-kinaesthetic perception. The task was carried out under three conditions of head position: (1) head upright, (2) head tilted 30° to the right, and (3) head tilted 30° to the left. The adjustments were carried out with: (a) only right hand, (b) only left hand, and (c) both hands simultaneously. Thus, the subjects were tested once in each of nine conditions.

The only difference between experiment 1 and 2 was the initial position of the rod. In experiment 1 this position, kept constant for all the items, was 60° to the left, in experiment 2 it was 60° to the right.

The number of the subjects in experiment 1 was sixty of each sex, and in experiment 2, thirty of each sex.

Experiment 3 dealt with visual perception. There were four conditions of head position: head tilted, (1) 30° to the right, (2) 30° to the left, (3) 45° to the right, and (4) 45° to the left. The position with head upright was omitted.

The subjects, thirty of each sex, were instructed to use the best hand for the adjustment of the rod. Each variable was tested four times; thus altogether sixteen measurements were made. The initial positions of the rod (60° to the left and 60° to the right) were rotated in the sequence ABBA BAAB and half the number of the subjects (men and women alike) had to begin with head tilted 30°, half the number with head tilted 45°.

III

RESULTS

The tactual-kinaesthetic experiments.

Tables I and II show the results of experiments 1 and 2. They confirm the general findings from the two studies referred to above: in tactual-kinaesthetic perception the rod is generally adjusted to the opposite side of the tilt of the head.

All the differences between right and left tilt of the head are fully significant. In *Table I*, CR computed from σ for the individual differences of the results between the

TABLE I

TACTUAL-KINAESTHETIC PERCEPTION WITH ROD INITIALLY SET 60° TO THE LEFT (EXPERIMENT 1). THE PLUS SIGN INDICATES HERE AND IN THE FOLLOWING TABLES A POSITION CLOCKWISE OF THE TRUE UPRIGHT

		Position of Head					
		Upright		30° Right		30° Left	
		M	σ	M	σ	M	σ
Women N = 60	Right hand.. ..	- 0.97	2.82	- 2.69	3.16	+ 0.45	3.19
	Left hand	- 1.57	3.16	- 3.20	3.87	+ 0.42	3.58
	Both hands.. ..	- 0.15	3.20	- 2.50	4.08	+ 1.83	4.13
Men N = 60	Right hand.. ..	- 1.08	2.95	- 2.08	2.37	- 0.17	2.99
	Left hand	- 0.90	2.40	- 1.70	3.55	+ 0.14	2.63
	Both hands.. ..	+ 0.01	2.55	- 1.06	3.09	+ 1.85	3.09

TABLE II

TACTUAL-KINAESTHETIC PERCEPTION WITH ROD INITIALLY SET 60° TO THE RIGHT
(EXPERIMENT 2)

		Position of Head					
		Upright		30° Right		30° Left	
		M	σ	M	σ	M	σ
Women N = 30	Right hand.. ..	- 0.79	2.41	- 3.59	1.74	+ 0.33	3.13
	Left hand	+ 0.42	2.83	- 3.15	3.81	+ 1.39	3.94
	Both hands.. ..	+ 0.53	2.94	- 2.45	4.39	+ 1.79	2.28
Men N = 30	Right hand.. ..	- 0.91	2.20	- 2.03	1.91	+ 0.58	2.75
	Left hand	- 0.23	3.02	- 1.19	4.07	+ 1.10	3.26
	Both hands.. ..	+ 0.43	2.83	- 0.85	3.51	+ 1.31	2.74

tilts of the head are as follows: Female subjects (the right hand, etc., in the sequence of the table), 8.24, 8.52 and 9.40; Male subjects, 6.62, 7.11 and 10.42.

The correlations between the right-tilt and left-tilt of the head are significantly positive for all hand variables. The coefficients (in the same order as above) for female subjects are 0.57, 0.63, and 0.62. Male subjects: 0.66, 0.83, and 0.75. Coefficients of consistency have earlier been reported in related experiments (Sandström, 1951, p. 88 ff.; Witkin, 1949 (*b*), p. 156 ff.). It is probably justifiable to consider the coefficients given here as particularly remarkable since they are based only on a single measure of each of the variables. Both our own earlier values and those obtained by Witkin are based on several measures of each item. Thus, a consistency in the performances seems to be established from the beginning of the trials. This may be of some interest for the further theoretical discussion.

The effect of the changed initial position of the rod is not great, as is shown in Table II. The average differences in results between the right-tilt and left-tilt of the head have been slightly increased for both sexes and the means of the results from upright position of the head are nearer the plumbline, but not in a significant way. Predominant deviations to the left, with head upright, are previously reported (Sandström, 1952), and this holds good also for visual perception (Neal, 1926).

If we consider the tables from the viewpoint of sex differences, we see that the differences in the results between the right and the left tilt of the head are greater throughout for the women. These differences were obtained by calculating the distance between the mean values of each of the three hand variables. For example for the right hand in Table I $(- 2.69^\circ) + (+ 0.45^\circ) = 3.14^\circ$; $(- 2.08^\circ) - (- 0.17^\circ) = 1.91^\circ$; $(3.14^\circ) - (1.91^\circ) = 1.23^\circ$.

	Right hand	Left hand	Both hands
Table I ..	1.23°	1.78°	1.42°
Table II ..	1.31°	2.25°	2.08°

The figures may be taken to indicate that the tactual-kinaesthetic phenomenon with which we are dealing here is more pronounced in the case of women. CR for

the differences in Table I are 2.52, 3.61, and 2.59, σ being directly computed from the individual differences between the two tilt variables under discussion.

Do these greater differences for women deviate equally to the left and to the right of the corresponding male values or is it possible to observe a deviation in any particular direction? We may point out the general predominance of the negative values also reported by Wapner and Werner (1952). Here this is shown very markedly in the female results. Only for the right hand in experiment 1 does the greater female difference deviate equally to the left and to the right (-0.61° and $+0.62^\circ = 1.23^\circ$). In all other cases the larger part of the difference falls on the negative side, and in two cases (both hands in experiment 1 and the right hand in experiment 2) both the compared value italicized figures below, fall to the left of the corresponding male results. The figures show how much the female subjects deviate to the left of the male subjects.

	Right hand	Left hand	Both hands
Table I ..	-0.61°	-1.50°	-1.44°
Table II ..	-1.56°	-1.96°	-1.60°

These figures may be compared with the preceding figures which showed the differences of the total amount of deviations between men and women. Women show significantly greater deviations to the left.

The visual experiment.

Table III shows the results of experiment 3 which is concerned with the visual perception of verticality.

TABLE III
VISUAL PERCEPTION (EXPERIMENT 3)

	Position of Head							
	30°				45°			
	Right		Left		Right		Left	
	<i>M</i>	<i>σ</i>	<i>M</i>	<i>σ</i>	<i>M</i>	<i>σ</i>	<i>M</i>	<i>σ</i>
Women N = 30 ..	-0.42	1.22	-1.41	1.62	-0.45	1.86	-1.27	2.15
Men N = 30 ..	-0.48	1.54	-1.19	1.84	-0.28	1.87	-1.04	2.22

The discussion of the A-phenomenon with small tilts of the head has been rather confusing. G. E. Müller (1916) called attention to the fact that with small head tilts the opposite of the A-phenomenon is often observed. This was called by Müller the E-phenomenon. (This is the effect generally obtained in tactual-kinaesthetic perception as has been shown both here and earlier.)

Wapner and Werner (1952) assert the generality of the E-phenomenon in visual perception as regards small tilts of the head (15° and 30°) and an earlier study (Werner, Wapner & Chandler, 1951) shows clearly significant deviations in the E-direction. Witkin and Asch interpret their results in the same way and state that there is a region of transition from the E- to the A-effect at about 42° of tilt (1948, p. 609). We have chosen our angles of tilt in accordance with this statement. In these studies the different stimuli were given by the experimenter in form of a step-wise procedure. "The E turned the rod slowly by increments of approximately 3° until S stated that

the rod was vertical." (Wapner, Werner & Chandler, 1951, p. 341 f.). It should be noted, that the average difference in our results between the two tilt variables is less than 3° . In a later study the procedure has been slightly changed. "The rod was moved in 5° steps up to the position of 10° (initially from 30°); then it was moved in 1° steps until 5° ; and finally, 0.5° steps were used until S reported that the rod was tilted to the opposite side. . . . Each judgment within a trial had to be made as quickly as possible: approximately 1 second was allowed." (Werner & Wapner, 1952, p. 68 f.). The purpose of the latter paper was to show the effect of the initial position on the apparent verticality. It was shown that the apparent upright is always closest to the position in which the rod was set initially for that trial; and this holds good also in our experiment (the shifts are in average about 0.5° , and alike for both sexes; these shifts within the tactual-kinaesthetic field—cf. Tables I and II—are less and not unambiguous).

There is no reason for a detailed discussion of this method here, but it may not be prudent to make direct comparisons of visual results obtained by a method of limits with tactual-kinaesthetic results obtained by a method of reproduction, as Werner and his co-workers have done. According to Witkin and Asch (1948) body tilt gives the same results as a corresponding head tilt in visual perception; and in tactual-kinaesthetic perception the study of Wapner and Werner (1952) shows that the effect of body tilt is similar to that of head tilt.

Mann (1951) reports results from small tilts of the head which are in keeping with the A-phenomenon, and to the extent that the present study gives tenable values for visual perception of verticality it supports Mann and also Müller.

From Table III it appears that the differences between the right and left tilt of the head are less throughout than in the tactual-kinaesthetic experiments. Actually we obtain only one fully significant CR; namely, if we put together men and women in the experiment with 30° of tilt. There we get $CR = 2.58$ (1 per cent. level). The CR for women separately is 2.25 (5 per cent.) and all other CRs are much below 2.00.

The small differences between the values for 30° and 45° do not support the statement of Witkin and Asch (1948). Actually our experiment shows more E-effects at 45° than at 30° . At 30° there are 43 A-effects and 17 E-effects; and at 45° the corresponding figures are 33 and 27 (no sex differences in the distribution). These figures in experiment 2 (tactual-kinaesthetic) are 3 and 57, and are thus not at all ambiguous.

The coefficients of correlation between the results from the right and left tilt of the head are all negative: 30° female subjects — 0.38, male subjects — 0.22; 45° female subjects — 0.62, male subjects — 0.35. In the tactual-kinaesthetic experiments these figures were all significantly positive.

If we consider Table III from the point of view of sex differences we cannot demonstrate any clear difference. There is also here, however, a slight tendency for women to deviate more to the left than is the case for men, but not in a significant way.

IV

DISCUSSION

Available studies on tactual-kinaesthetic perception of verticality with head tilted sidewise show unambiguous results. This is not true for the corresponding visual studies. Our study may support the original statement of Aubert and also Müller's opinion that a phenomenon contrary to the A-phenomenon appears often but not generally, at small tilts of the head. Irrespective of the kind of perception the minus deviations dominate.

Calling attention to this general predominance for adjustments to the left, we define the two types of phenomenon as a relation between the two apparent uprights as determined under different conditions of sidewise head tilt. Whether these uprights fall into the same quadrant or not seems to be a secondary matter.

Concerning sex differences, we have shown in this paper that the deviations are more pronounced for women than for men in tactual-kinaesthetic perception and, moreover, that these significantly greater deviations fall to the left of the male values. They do not fall equally to the right and to the left of the corresponding male values as might have been expected. No significant sex differences could be found within the visual field of perception.

Phenomenologically, the subjects experienced the tactual-kinaesthetic task as much more difficult than the visual task. In the latter case scarcely anyone found the adjustment to the upright difficult in any way, but the tactual-kinaesthetic situation gave rise to much uncertainty.

In spite of this fact, however, the correlations between the tilt-variables are significantly positive in the tactual-kinaesthetic experiments and negative throughout in the visual trials. Evidently we have a much better experience of the visual than of the tactual-kinaesthetic vertical. This seems to have the effect of reducing the general deviations as well as the sex differences. In view of the above, it seems strange that the consistency of visual perceptions is less than for tactual-kinaesthetic perceptions. How to interpret the negative correlations in an adequate way is an open question. Perhaps these effects (and the general predominance of adjustments to the left) can explain to some extent the discrepancies of the results and clarify the discussions concerning Aubert's phenomenon when found with small tilts of the head.

On the basis of clear sex differences obtained from several experiments on verticality Witkin (1949 (a), 1949 (b)) has built up a theory of sex difference. Witkin says that his findings "indicate that women, in judging the upright, tend to be more dependent than men on the surrounding field and to rely less on experience from their own bodies" (1949 (a), p. 24). His tests "involve a conflict between the main (visual and postural) determinants of the perceived upright. There each presents the problem of integrating these conflicting sensory experiences and arriving at a unified perceptual impression." (1949 (b), p. 151). Thus every result represents a compromise which is the outcome of this conflict. This theory is very simple and attractive, and perhaps explains to a certain extent the results reported in the present paper. It seems difficult, however, to accept the relation between a greater conflict and an increased consistency as has been shown here.

In an earlier study (1953) we obtained sex differences in the opto-kinaesthetic field. These very marked differences do not seem to fit into the Witkin theory, mainly because of the absence of conflicting sensory experiences in these situations. There were no changes of the postural determinant (no body or head tilts) in these experiments.

In a recent paper (1954 a) published in Swedish, we have taken up this question starting from some critical comments on the Witkin theory. We have not been able to formulate a theory that also satisfies our opto-kinaesthetic findings; but it is our intention to make a comprehensive survey of the relations between our findings and the Witkin theory, thereby hoping to elucidate this theoretical dilemma.

REFERENCES

- AUBERT, H. (1861). Eine scheinbare bedeutende Drehung von Objecten bei Neigung des Kopfes nach rechts oder links. *Virchows Arch.*, **20**, 381-393.
 MANN, C. W. (1951). The effects of auditory-vestibular nerve pathology on space perception. *J. exp. Psychol.*, **42**, 450-456.

- MÜLLER, G. E. (1916). Ueber das Aubertsche Phänomen. *Z. Psychol.*, **49** (II), 109-246.
- NEAL, E. (1926). Visual localization of the vertical. *Amer. J. Psychol.*, **37**, 287-291.
- SANDSTRÖM, C. I. (1951). *Orientation in the present space*. Stockholm.
- SANDSTRÖM, C. I. (1952). Taktil-kinestetisk bestämning av lodlinjen med sidolutat huvud (Tactual-kinaesthetic determination of the vertical with head tilted). *Nord. Psykol.*, **4**, 156-165 (English summary).
- SANDSTRÖM, C. I. (1953). Sex differences in localization and orientation. *Acta Psychol.*, **9**, 82-96.
- SANDSTRÖM, C. I. (1954). Könsskillnader vid taktil-kinestetisk lodbestämning (Sex differences in tactual-kinaesthetic determinations of the vertical). 3. *nord. psykol. mötet i Helsingfors*, 1953. (3rd Scand. Congr. Psychol. Helsinki, 1953.)
- WAPNER, S., and WERNER, H. (1952). Experiments on sensory-tonic field theory of perception: V. Effect of body status on the kinaesthetic perception of verticality. *J. exp. Psychol.*, **44**, 126-131.
- WAPNER, S., WERNER, H., and CHANDLER, K. A. (1951). Experiments on sensory-tonic field theory of perception: I. Effect of extraneous stimulation on the visual perception of verticality. *J. exp. Psychol.*, **42**, 341-345.
- WERNER, H., and WAPNER, S. (1952). Experiments on sensory-tonic field theory of perception: IV. Effect of initial position of a rod on apparent verticality. *J. exp. Psychol.*, **43**, 68-74.
- WERNER, H., WAPNER, S., and CHANDLER, K. A. (1951). Experiments on sensory-tonic field theory of perception: II. Effect of supported and unsupported tilt of the body on the visual perception of verticality. *J. exp. Psychol.*, **42**, 346-350.
- WITKIN, H. A. (1949a). Sex differences in perception. *Trans. N.Y. Acad. Sci.*, **12**, 22-26.
- WITKIN, H. A. (1949b). The nature and importance of individual differences in perception. *J. Pers.*, **18**, 145-170.
- WITKIN, H. A., et al. (1954). *Personality through perception*. New York.
- WITKIN, H. A., and ASCH, S. E. (1948). Studies in space orientation. III: Perception of the upright in the absence of a visual field. *J. exp. Psychol.*, **38**, 603-614.

FURTHER EXPERIMENTS ON THE OLFACTORY THRESHOLDS OF PURE CHEMICAL SUBSTANCES, USING THE "SNIFF-BOTTLE METHOD"

BY

G. H. CHEESMAN and M. J. TOWNSEND

From the Departments of Chemistry and Psychology, Reading University

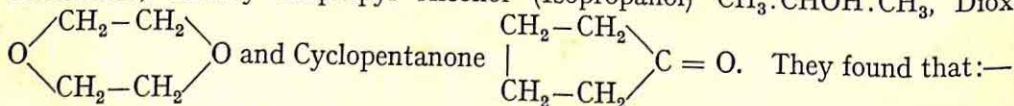
The "Sniff-Bottle" method of determining olfactory thresholds under conditions of adaptation to a second olfactory stimulus, as described by Cheesman and Mayne (1953), has been further developed. Their study was limited to the mutual adaptation effects of three substances. The original matrix of nine quantities, described as degrees of community of odour property, has been revised and extended to include a fourth stimulus-substance. It is suggested that the matrix now obtained shows a definite pattern characteristic of consistent mutual relationships between the substances. A number of shortcomings in the original technique are discussed, and the appropriate emendations described.

I

INTRODUCTION

Cheesman and Mayne, in experiments reported in this *Journal* in 1953, attempted to establish a method, the "sniff-bottle group test," for the determination of olfactory thresholds. They applied the technique to the investigation of changes in the thresholds of pure chemical substances occasioned by adaptation of the olfactory mechanism by the same, or a different substance. One of their objects was to ascertain whether there exists any "community of odour-property" (C.O.P.) between different substances, which might manifest itself in ordered mutual effects, and might be related to factors of chemical constitution. They write ". . . when the values of the logarithm of the threshold concentration of the test stimulus are plotted against the logarithms of the concentration of the masking stimulus for a pair of odorous materials, a straight line is obtained," and suggest that the gradient of this line may be used as a measure of the community of odour-property between the two substances.

Using this basic method, they investigated the properties of three water-soluble substances, namely Isopropyl Alcohol (Isopropanol) $\text{CH}_3\cdot\text{CHOH}\cdot\text{CH}_3$, Dioxan



1. The straight-line relationship, or a close approximation to it, subsists over at least a part of the adaptation range.
2. In each case where the adapting and test substances were the same (i.e. in "homogeneous" tests) the slope was approximately 0.7.
3. All nine values in the 3×3 matrix, representing all possible combinations of adapting and test stimuli, lie between + 0.2 and + 1.0. (Cheesman and Mayne erroneously attributed a negative sign.)

The present paper describes experiments designed to improve upon Cheesman and Mayne's technique, to correct some of their determinations, and to extend the matrix

to include a fourth stimulus-substance, namely Cyclopentanol $\begin{array}{c} \text{CH}_2-\text{CH}_2 \\ | \\ \text{CH}_2-\text{CH}_2 \end{array} \text{CHOH}$.

II

REVISION OF EXPERIMENTAL PROCEDURE

Our experience in repeating some of the tests carried out by Cheesman and Mayne suggested that their procedure requires some revision and standardization if repeatable results are to be obtained. It should be noted that Cheesman and Mayne were concerned not only to obtain comparable threshold values, but also to ascertain the effects of different experimental procedures. The following aspects especially need attention:—

1. *The range of concentration of the test-stimulus.* It will be recalled that Cheesman and Mayne usually employed nine test-bottles, together with three "dummies" containing no test-substance. (We shall refer to the *strongest* bottle as the *first* bottle, and the *weakest* as the *ninth*). Generally the range from the strongest to the weakest stimulus was five units on a logarithm-to-the-base-two scale. In some experiments, however, the range was increased to seven or ten units. The threshold-values so obtained were commonly among those which deviated most from the straight-line fitted to the plot of threshold against adapting stimulus-strength by the method of least squares. In all new determinations reported in this paper, therefore, the bottle-series covered five $\log_2$ steps, using nine bottles each differing from the next by half a $\log_2$ unit.

2. *Position of the threshold within the test-stimulus range.* In experiments such as these, where a considerable threshold shift from adaptation is in question, it is important, of course, to try to arrange the actual series of stimulus-strengths in any given determination so that the threshold comes well within the stimulus-range. This is not always easy to manage, since there is in practice no simple way of getting a rough preliminary estimate of the threshold. In fact, in Cheesman and Mayne's experiments, a number of cases occurred in which the threshold lay well towards one end or the other of the bottle-series. As would be expected, the values in these cases deviated from what might be expected from a straight-line relationship in an upward direction if there are too many supraliminal bottles, and in a downward if there are too many subliminal. If the threshold lies towards the strong end of the series—i.e. there are too many subliminal bottles—it becomes questionable whether the subjects know what they are looking for in the test bottles. Such uncertainty is detrimental to their consistency of judgment. They are therefore presented with a sample bottle at the outset of the same strength as the strongest of the series. If they are still in any doubt as to the presence or character of the odour, a still stronger bottle is offered.

If, on the other hand, too many supraliminal bottles are included in the series, the threshold may be displaced, or a marked degree of heterogeneity introduced in to the results. Inspection of the individual responses suggests strongly that these effects are due to 30 seconds being an insufficient interval to allow between successive test-sniffs. There is not full recovery from adaptation to the preceding test-stimulus. Relatively strong stimuli following strong stimuli are commonly recorded as minus while, by comparison, relatively weak stimuli preceded by weak stimuli tend to be recorded as plus. Not only is the threshold displaced in an upward direction, but a significant χ^2 often emerges when the points are fitted to a straight line in the course of probit analysis.

3. *Positive responses to no-stimulus bottles.* In any experiment in which *vexiersuchen* are introduced, a certain number of positive responses to non-existent stimuli will be obtained. In work such as the present, however, where a *group* of subjects is used, the presence of such responses may be due to either or both of two factors. In the first place, it may indicate unreliability of individual subjects. Cheesman and Mayne in fact rejected the data from any subject who gave positive responses to more than one of the three no-stimulus bottles. But in the second place, the total number of such responses is an indication of the general reliability of the experiment as a whole. A high proportion may be occasioned by an unsettled attitude in the group, or by contamination of the bottles, for example. An experiment in which more than 20 per cent. of the total responses to the no-stimulus bottles are positive should be rejected as a whole and, reviewing the circumstances, the experimenter is often able to make a guess at a possible cause for the failure. Since the "unreliable" subjects, with more than one positive response, contribute equally with the "reliable" to the total number of positive responses, which is treated as an index of reliability of the experiment as a whole, we have found it important to assess the reliability of each experiment *before* discarding the unreliable subjects. The total number

of positive responses to no-stimulus bottles is first counted, and if this is greater than 20 per cent. the experiment is rejected *in toto*. If not, those subjects having more than one positive response are discarded, and a threshold value calculated from the amended figures.

In the statistical treatment of their data by probit analysis, Cheesman and Mayne took account of the positive responses to no-stimulus bottles by applying Abbott's correction for "natural mortality" to these cases (Finney, 1952). The appropriateness of this procedure may be questioned, but it should be noted that when it is carried out, the tendency of Abbott's correction is to produce an artificial raising of the threshold which is the greater the higher the proportion of positive responses.

4. *Number of Subjects in a group.* In some of the earlier experiments of Cheesman and Mayne, the responses of as few as 8, or as many as 34, subjects were used for treatment by probit analysis. Mayne (1953) had already established that if the number of subjects (n) was greater than 20, the shorter method of obtaining fiducial limits may be applied (Finney, 1952, p. 63). Latterly, however, it has been noted by the present authors that if n is much greater than 20, a noticeable weakening of concentrations and some intrusion of foreign odours was liable to occur. In our new determinations we have, therefore, adhered to groups of approximately 20 subjects, and have rejected those determinations of Cheesman and Mayne to which this criterion does not apply.

5. *Cleaning of Bottles.* The method of cleaning of bottles used by Cheesman and Mayne consisted in:—

- (1) Washing in concentrated sulphuric acid.
- (2) Rinsing thoroughly in cold and hot water.
- (3) Sterilizing by boiling in distilled water for three-quarters of an hour on three consecutive days.

The use of sulphuric acid, however, has been discontinued in favour of chromic nitric, or chromic sulphuric acid mixture made by dissolving 40 gms. of Potassium Dichromate in a litre of concentrated sulphuric acid and adding 100 ml. water. In addition to being a better cleansing agent in view of its oxidising power, this has the further advantages that bottles are less likely to break, and that splashes of the cleansing material can easily be detected by its colour. The bottles, after washing in the mixture, are left stoppered and upside down for at least two hours. This is followed as before by boiling and rinsing. This procedure has been associated with a lowered percentage of positive responses to no-stimulus bottles.

6. *Instructions given to subjects.* New instructions to the subjects have been introduced by a new experimenter, and are therefore included:—

"The little bottles contain solutions of . . . all at *different* concentrations and what I want to do is to find the least concentration of these that you can detect immediately after sniffing from the big bottle which is on your table. The procedure is as follows (demonstrate)—take one good sniff from the big bottle, breathe out completely through your nose and then take a good sniff from the little bottle which will be on your table as well. All you have to do is to record the letter on the top of the little bottle, a plus if you detect the odour, a minus if you can't, or you may record $\pm$ if you are not sure, but try and make up your mind. This is a trial bottle, the same concentration as the strongest bottle, which goes round before the test so that you will be familiar with the procedure and *also* so that you will know the smell you are looking for in each of the little bottles. We have to allow 30 seconds between each bottle so that your nose has time to recover so that I will give you the signal when to sniff from the big bottle, go on to the little one at your own time and record your result. There are two precautions to be taken—if you are not sure about the little bottle, please don't take any second sniffs—if you do the effect of the big bottle will have gone off, and this is what we are investigating—if you really can't make up your mind record $\pm$. Secondly will you shake the big bottles between tests—we will look after the little ones as we pass them on."

As the subjects are given the first signal, they are reminded by the words "big bottle," "little bottle," "record your result and wait for the next signal."

We may briefly summarize the modifications of technique, and the criteria for rejecting some previous determinations by Cheesman and Mayne.

- (1) Range of concentration of test-stimulus standardized—nine stimuli in steps of half $\log_2$ units.
Cheesman and Mayne's determinations rejected if total range of concentration greater than five $\log_2$ units.
- (2) Sample bottle containing stimulus equal in strength to strongest test-bottle given to subjects. If necessary an even stronger bottle offered.
- (3) Determination rejected if threshold falls below the eighth bottle, and/or if χ^2 is significant on fitting probits.
- (4) Determination rejected if threshold falls above the fifth bottle.
- (5) Experiment rejected as a whole if more than 20 per cent. positive responses to no-stimulus bottles.
Subjects rejected if more than one positive response to no-stimulus bottles.
- (6) New bottle-cleaning procedure.
- (7) Number of subjects in group held constant at 20–22.
Cheesman and Mayne's determinations rejected if $n > 26$.
- (8) New instructions.

The adoption of these more rigorous experimental standards has required the rejection of quite a high proportion of Cheesman and Mayne's values for the community of odour-property. It was decided not to repeat the experiments using homogeneous pairs of substances, but the other values have all been redetermined by the present authors. All determinations involving the new, fourth, stimulus substance have been carried in accordance with the new standards. In all other respects our procedure has been that described in Cheesman and Mayne (1953).

III

TREATMENT OF DATA

Certain anomalous features in the application of probit analysis to data obtained in experiments such as these were noted by Cheesman and Mayne. These relate especially to the estimation of fiducial limits, and their effect is to set these limits at artificially low values. While no exception can be taken to the use of probit analysis for the computation of a threshold, the use of the fiducial limits afforded is unjustifiable. And it cannot be denied that the method is clumsy and time-consuming if used merely for the former purpose.

Unfortunately mathematical difficulties in the development of a means of estimating correctly the fiducial limits appear to be at present insuperable. Another simpler method of arriving at a threshold-value and an expression of its reliability is as follows:—The subjects' responses are arranged in order of concentrations, and each subject is allotted a threshold value—the individual threshold (I.T.) such that the number of *minus* responses above the threshold equals the number of *plus* responses below. The *plus-minus* responses are counted as half, or if more convenient the I.T. is taken as the midpoint between the two values counting the *plus-minus* first as a *plus* and then as a *minus*.
E.g. Conc. (x -scale):

1.0	+	+	+	+	+	—	+
1.5	+	+	+	+	+	—	+
2.0	+	+	+	+	+	—	+
2.5	—	±	±	±	—	—	+
3.0	+	+	±	—	—	—	+
3.5	+	—	—	—	+	—	+
4.0	—	—	—	—	+	—	+
4.5	—	—	—	—	+	—	+
5.0	—	—	—	—	+	—	+
I.T.	3.25	3.00	2.75	2.50	4.25	0.75	5.25

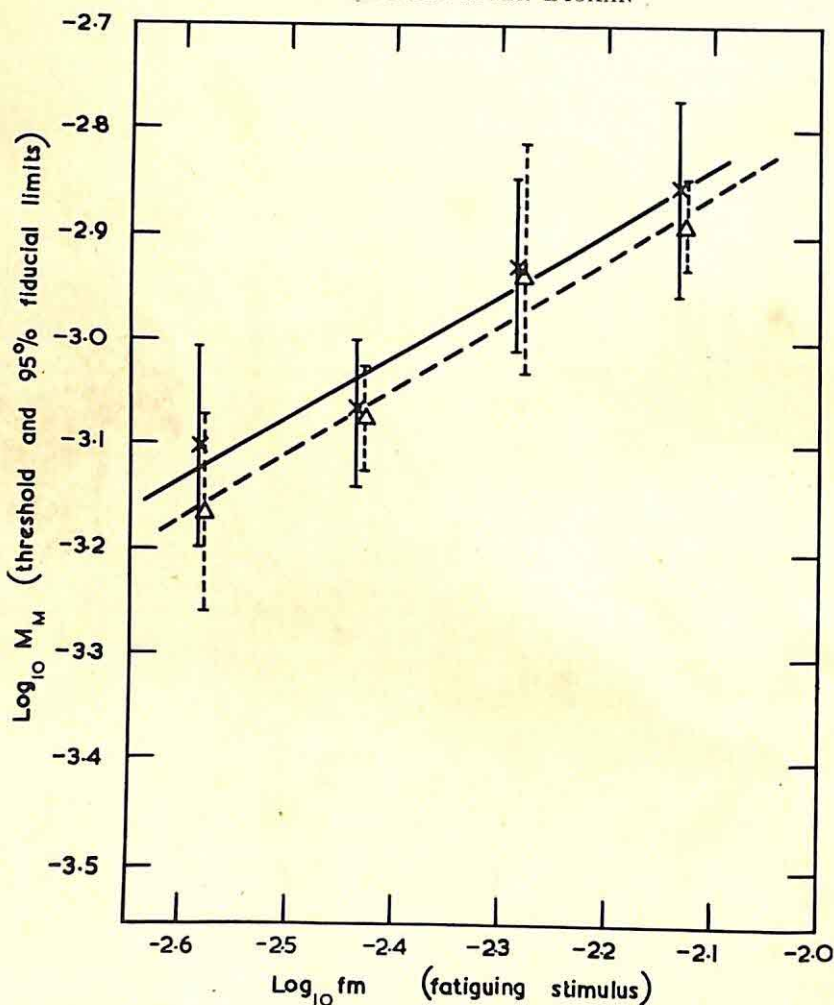
The Mean Threshold (M.T.) is calculated = $\frac{\Sigma(I.T.)}{N}$ where N is the number of subjects

$$\text{The Standard Error} = \sqrt{\frac{\Sigma(I.T.)^2 - (M.T.)^2.N}{N(N-1)}}$$

The 95 per cent. fiducial limits = $(M.T.) \pm t \times (S.E.)$, t being read from the tables with N degrees of freedom.

To allow comparison with the results of Cheesman and Mayne, our own determinations, and the corresponding figures for "community of odour-property" have been computed by both probit analysis and I.T. method. Figure 1 gives a specimen plot from which the relationship of the two values can be seen.

FIGURE 1
ISOPROPYL ALCOHOL AFTER DIOXAN



Solid lines—results analysed by probit analysis
Broken lines—results analysed by individual threshold method.

IV

RESULTS

It will be recalled that Cheesman and Mayne's procedure was for each pair of substances, plot, $\log_{10}$ (threshold) against $\log_{10}$ (concentration of adapting stimulus). A straight line is then fitted to the points by the method of least squares, using

$1/(S.E.)^2$ as a weighting factor. The slope of this line is then measured and taken as a measure of the degree of community of odour property. The values resulting from our determinations, together with those of Cheesman and Mayne which we regarded as satisfactory, are shown, for the two methods of threshold computation, in Tables 1 and 2. The numbers in brackets adjacent to each figure show the number of points established and used in fitting the straight line.

TABLE I
COMMUNITIES OF OLFACTORY PROPERTY (PROBIT ANALYSIS METHOD)

		<i>Test Stimulus</i>			
		<i>Isopropyl Alcohol</i>	<i>Dioxan</i>	<i>Cyclopentanone</i>	<i>Cyclopentanol</i>
<i>Adapting Stimulus</i>	<i>Isopropyl Alcohol</i>	0.68 (5)	0.51 (3)	0.24 (3)	0.72 (3)
	<i>Dioxan</i>	0.59 (4)	0.74 (5)	0.48 (3)	0.50 (5)
	<i>Cyclopentanone</i>	0.18 (4)	0.54 (4)	0.67 (9)	0.04 (3)
	<i>Cyclopentanol</i>	0.64 (3)	0.59 (4)	0.17 (4)	0.68 (3)

TABLE II
COMMUNITIES OF OLFACTORY PROPERTY (INDIVIDUAL THRESHOLD METHOD)

		<i>Test Stimulus</i>			
		<i>Isopropyl Alcohol</i>	<i>Dioxan</i>	<i>Cyclopentanone</i>	<i>Cyclopentanol</i>
<i>Adapting Stimulus</i>	<i>Isopropyl Alcohol</i>	0.67 (5)	0.47 (3)	0.19 (3)	0.70 (3)
	<i>Dioxan</i>	0.61 (4)	0.88 (5)	0.47 (3)	0.57 (5)
	<i>Cyclopentanone</i>	0.16 (4)	0.56 (4)	0.62 (9)	0.07 (3)
	<i>Cyclopentanol</i>	0.66 (4)	0.65 (4)	0.19 (4)	0.66 (3)

V

DISCUSSION

From the general conditions of the experiment, and the method of handling the data, it seems clear that no very great significance attaches to the second decimal place in these values, and this is borne out by the differences which arise as a result of

computing the threshold in two different ways. No attempt at a detailed analysis of the matrix would be justifiable, but the following provisional conclusions, not without interest, may be drawn:—

1. All homogeneous pairs (i.e. test and adapting stimulus the same) have the same C.O.P., namely 0.7 approximately. It should be noted that this is a genuine finding, there being no *a priori* reason why this should be the case, nor any feature of the experimental design which makes it so.
2. Reciprocal pairs of substances (i.e. test and adapting stimuli reversed) have approximately the same C.O.P.
3. No C.O.P. is greater than that for homogeneous pairs.
4. The two alcohols, isopropanol and cyclopentanol have the same effect as adapting stimuli on each other as each does upon itself. Moreover both have the same C.O.P.'s with the other substances. On the other hand, the two compounds with the same radical, namely cyclopentanone and cyclopentanol have markedly different C.O.P.'s. with other substances.

We do not at present propose to offer any theoretical interpretation of these findings.

Chemical purity and olfactory properties. In the case of one substance, isopropanol, it was necessary in the course of our redeterminations of Cheesman and Mayne's threshold values to start using a fresh bottle of the material. Both bottles were of analytical reagent standard, and had come from the same manufacturer. Of the stated maximum impurities, only ketones and aldehydes (0.03 per cent.) could be supposed to be relevant to the olfactory properties of the sample. Yet the redetermined thresholds were found to be markedly displaced. When, for instance, some of Cheesman and Mayne's values for isopropanol with cyclopentanone as adapting stimulus were being redetermined, displacements of up to $1.5 \log_{10}$ units were found, and in consequence the rest of the values for this pair of stimuli not previously rejected on account of our more rigid standards, were also checked. It was then seen that the amount of displacement was approximately constant, and that the new straight line lay roughly parallel to the first. Since only the gradient of the line is involved in the final results, it would seem that effect of using more than one sample was not serious. But the observation arouses some interesting reflections regarding the effects of impurity upon the olfactory properties of substances. The difference in amount of impurity between the two samples cannot have been more than three parts in 10,000, yet its effect appears to have been to produce a 400 per cent. shift in threshold. It would seem that chemical purity is an important problem in olfactory research. Experiments relating to it are in progress.

The authors wish to express their thanks to Professor R. C. Oldfield for many helpful discussions and his continued interest in this work, to Miss V. R. Cane for her assistance, and to Uniliver Limited for the financial support which has made this work possible.

REFERENCES

- BURSILL, A. E. (1951). A comparison of two methods for obtaining the median olfactory thresholds of a sample population. B.A. Thesis, Reading University.
 CHEESMAN, G. H. and MAYNE, S. (1953). The influence of adaptation on absolute threshold measurements for olfactory stimuli. *Quart. J. exp. Psychol.*, **5**, 22–30.
 FINNEY, D. J. (1952). *Probit Analysis*. Cambridge.
 MAYNE, S. (1953). Determination of olfactory thresholds for pure chemical substances. Ph.D. Thesis, Reading University.

SOME EFFECTS OF SEVERE HYPOTHERMIA ON LEARNING AND RETENTION

BY

R. K. ANDJUS¹, F. KNÖPFELMACHER², R. W. RUSSELL
and AUDREY U. SMITH

*From the National Institute for Medical Research, Mill Hill, London, and the
Department of Psychology, University College, London.*

Two experiments have been conducted to assess the effects of extreme hypothermia on learning and retention. In the first experiment two groups of rats were subjected to identical treatments except that the deep body temperature of one group was lowered to between $+1^{\circ}$ and 0° C. before both groups began learning a serial problem solving task. The second experiment included three groups of animals: a control and two experimental groups, the deep body temperature of animals in one of the latter being reduced to between $+13.4^{\circ}$ and $+18.5^{\circ}$ C. and that of the other to between $+1^{\circ}$ and 0° C. Except for the "hypothermia" phase of the design all groups in the second experiment received identical treatments in the following order: Training on a maze, hypothermia, tests for retention of the maze habit, training on the serial-problem-solving task, tests for speed of locomotion.

Animals whose body temperature had been lowered to between $+1^{\circ}$ and 0° C. showed a significant impairment in learning performance when compared with control animals. On the other hand, no significant differences were found between the performances of control animals and those cooled to between $+13.4^{\circ}$ and $+18.5^{\circ}$ C. The results also suggest that the effects of severe hypothermia may be temporary, since the percentage of impairment decreased as the interval between body cooling and learning increased.

The differences in retention of the maze habit among experimental and control groups were very small and in no instance were they statistically significant, although there was a consistent trend towards poorer retention following hypothermia. These small differences may be functions of the technique used to reduce deep body temperature rather than of the effects of hypothermia *per se*.

These results are based upon observations of the behaviour of non-hibernating, homoiothermic animals. With such animals, extreme hypothermia, such as that employed in the present study, results in complete arrest of heart beat, circulation, and respiration. It also suppresses electrical activity in the brain and in our animals cerebral activity may have been arrested for as long as $1\frac{1}{2}$ to 2 hours. Within the limits of our experimental procedure we have failed to find evidence that arrest of these vital metabolic processes as a result of hypothermia produces any very serious, permanent effects on the animal's behaviour once it has been successfully reanimated. If, as previous writers have suggested, severe hypothermia can be used to "... stop all nerve impulses in the brain momentarily ..." (Gerard, 1953), our results are difficult to explain if long-term memory is dependent upon the continuous activity of the brain.

I

INTRODUCTION

Recent developments in techniques for lowering the deep body temperature of homoiothermic animals have led to an increased interest in the effects of body temperature on physiological processes. Research has shown that, as deep body temperature decreases, such processes as oxygen consumption, pulse rate (Adolph 1950), and arterial pressure (Crismon, 1944) at first increase and then decrease, in many

¹ Of the Faculty of Science, University of Belgrade.

² Of the Department of Psychology, University of Melbourne.

cases approaching zero as the temperature reaches $+15^{\circ}\text{C}$. Other processes, such as heart rate (Crismon, 1944), conduction of cardiac impulses and cerebral metabolism (Fazekas and Himwich, 1939), decrease progressively as the body cools. A crucial point appears at about $+15^{\circ}\text{C}$., below which arrest of heart beat and respiration occur in the majority of non-hibernating mammals so far studied (Burton and Edholm, 1955). Until very recently it was generally believed that this was the lowest level of hypothermia compatible with their survival.

However, in 1951, Andjus reported that adult rats had been reanimated after cooling to much lower body temperatures and after prolonged respiratory and cardiac arrest had occurred. Since then improvements in the technique have been described (Andjus and Smith, 1954; Andjus, 1955; Andjus and Lovelock, 1955; Andjus and Smith, 1955) for reanimating rats subjected to extreme hypothermia and detailed information has been reported regarding a number of physiological changes occurring during the cooling and reanimation procedures. The first stage in this procedure involved narcotizing the animals by anoxia and hypercapnia as well as by exposure to cold so that they become flaccid and lethargic with colonic temperatures at or below $+20^{\circ}\text{C}$. Animals then cooled to body temperatures as low as $+15^{\circ}\text{C}$. re-warmed spontaneously when left at ambient temperature of $+20^{\circ}\text{C}$. and appeared to suffer no after-effects despite the severe treatment to which they had been exposed. Other animals were cooled in melting ice until their body temperatures were lowered to between 0° and $+1^{\circ}\text{C}$. These animals recovered completely even after 60 to 75 minutes at body temperatures below $+15^{\circ}\text{C}$., and after 40 to 60 minutes without detectable heart beat, circulation, or respiration. Their skulls as well as the rest of their bodies had been packed in ice during this same period and could have been little, if at all, warmer than the colon from which deep body temperature had been recorded. Despite this extreme cooling Andjus and Smith (1955) report no evidence of the neurological damage to the brain described by other researchers (Jensen and Parkins, 1954). There is evidence that electrical activity in the rat brain is suppressed at body temperatures below $+18^{\circ}\text{C}$. (Horsten, 1949; Lemaitre, 1954). Using this evidence, Andjus and Smith (1955) have suggested that, in their animals, "... cerebral activity was probably arrested for $1\frac{1}{2}$ to 2 hours." Animals exposed to this severe treatment recovered reflexes and posture within two hours of re-warming and showed apparently normal activity and behaviour within two or three days.

To the physiological psychologist, interested in relations between behaviour and the metabolism of the body as a whole or of various of its systems, the new techniques for inducing severe hypothermia offer fresh opportunities for research. The general idea is not a new one, although we have not found any previous study which reports the effects of complete cardiac and respiratory arrest on the behaviour of homoiothermic animals. In 1932 Hunter reported on the effect of inactivity produced by cold upon learning and retention in the cockroach, *Blatella germanica*. His hypothesis was "... that almost perfect retention might be secured if the subject's physiological processes could be held practically in quiescence during the period of rest interpolated between learning and relearning." The learning and retention of various groups of subjects exposed to temperatures of 3° to 6°C . were compared with the learning and retention of control groups. The results were contrary to Hunter's hypothesis, for both learning and retention of a simple darkness-avoiding response were poorer in the experimental than in the control groups. Hunter suggested the possibility that these results might be accounted for in terms of physiological changes occurring below a critical temperature of about 10°C .

More recently Gerard (1953) has suggested the use of hypothermia as "... a

simple way to test the question as to whether a memory is purely dynamic. One need only stop all nerve impulses in the brain momentarily and observe whether a memory is lost." The use of hypothermia for this purpose is illustrated in a study reported by Ransmeier (1953). Hamsters, which are hibernating mammals, were used as subjects. Deep body temperatures of animals anaesthetized by nembutal were reduced to approximately $+ 12^{\circ}\text{C}$. in certain groups and to approximately $+ 5^{\circ}\text{C}$. in others, and the effects of these treatments on behaviour in a maze were studied. The maze contained seven choice points spaced at regular intervals over the 6-foot length of a 9-inch-wide board. Its odd-even reliability is reported to be high, about 90 per cent. Two research procedures were used, the first consisting of the following order of events: three consecutive maze trials, hypothermia, two days rest, three more trials, etc. The second procedure was the same as the first except that hypothermia was induced after each trial, one trial being given every two days. The results indicated that, when the hypothermia procedure was begun at *one minute* after completion of a learning trial, significant impairment of performance on subsequent runs occurred. However, when the procedure began *five minutes* after a trial, no significant effects on later performance were found. From these interesting results Ransmeier postulated that "Normal retention events . . . include an initial conducted activity stage (reverberations?) and resultant enduring material changes (size of neuron ends?)" He was also careful to point out that ". . . almost *all* of the procedures involving onset of marked activity or injection of any sort *at one minute* after a learning run, gave increases over normal scores" The impairment of later performance following induction of severe hypothermia at one minute after a learning trial may have been due to certain aspects of the technique rather than to reduced body temperature *per se*.

These studies of Hunter and Ransmeier were concerned with the behaviour of poikilothermic animals and hibernating mammals which, under natural conditions, tolerate reduction of body temperature to levels lethal to man and most warm-blooded animals. Techniques (Andjus and Lovelock, 1955; Andjus and Smith, 1955) used in the present experiments make it possible to lower drastically the deep body temperature of non-hibernating, warm-blooded animals with very high percentages of successful recovery. Two or three days after such treatment it is difficult, if not impossible, to differentiate between control animals and animals whose deep body temperature has been lowered as far as $+ 1^{\circ}$ to 0°C . on the basis of their behaviour when living in adjacent cages.

The effects, if any, of hypothermia on behaviour are of interest to medical research workers as well as to students of animal behaviour. Hypothermia is a useful form of anesthesia for human surgery and is being used more and more frequently. It is important to know as much as possible about any secondary effects which may arise, such as effects on behaviour.

The present research was designed to study the effects on learning and retention of reducing the deep body temperature of non-hibernating, homoiothermic animals. Rats were used as subjects, different groups being cooled to different levels: $+ 13.4^{\circ}$ to $+ 18.5^{\circ}\text{C}$., the lowest temperatures consistent with spontaneous recovery, and $+ 1^{\circ}$ to 0°C . As has been pointed out above, very drastic changes in physiological processes occur as body temperature is lowered to $+ 15^{\circ}\text{C}$., and below this crucial point heart beat, respiration and electroencephalographic signs of brain activity disappear. If long-term retention is dependent upon the continuous normal activity of the central nervous system, it would be expected that tests for retention following these experimental treatments would show impairment. It might also be expected that, as with other drastic treatments such as electrical stimulation severe enough

to induce convulsions (Braun, Russell and Patton, 1949), learning of new habits would be impaired, at least temporarily.

II

EXPERIMENTAL DESIGN AND PROCEDURE

Two experiments have been conducted.

Experiment 1. This experiment was designed to determine the effects of reducing deep body temperature to $+1^{\circ}$ to 0° C. on two aspects of behaviour – speed of locomotion and serial problem solving. The experimental design is shown in Table I.

TABLE I
EXPERIMENT 1. EXPERIMENTAL DESIGN

Group	Phase			
	1	2	3	4
ES	Hypothermia: $+1^{\circ}$ to 0° C.	Preliminary Training	Serial Problem Solving	Speed of Locomotion
C	No hypothermia	Preliminary Training	Serial Problem Solving	Speed of Locomotion

Treatment of the two groups of animals differed only in Phase 1, when the deep body temperature of the ES animals was reduced to $+1^{\circ}$ to 0° C.; the C group animals served as controls. Each group consisted of eight male hooded rats from the animal colony of the National Institute for Medical Research. They were approximately 70 days old at the beginning of the experiment and were assigned at random to groups ES and C.

Phase 1 of the experiment was carried out at the National Institute and the technique described by Andjus and Smith (1954) was used. Cooling took place in two stages, body temperature being reduced first to approximately 15° C., then to between $+1^{\circ}$ and 0° C. About 20 minutes after the second stage began no more heart beats could be recorded and colonic temperature fell to approximately $+1^{\circ}$ over a further period of about 40 minutes. Reanimation consisted essentially in re-establishing the circulation by applying heat locally to the cardiac area before re-warming the whole body. This was accompanied by positive-pressure artificial respiration. The first heart beats were seen before the colonic temperature reached $+5^{\circ}$ C. and these were followed, in most animals, by increases in heart rate with continued re-warming. Survivors had to be kept in a warm atmosphere of $+28^{\circ}$ to $+32^{\circ}$ C. for two or three days after reanimation. Using the improved techniques of Andjus and Lovelock (1955), approximately 75 per cent. of the animals recovered from this treatment.

Between 19 and 21 days after reanimation preliminary training for the tests of serial problem solving behaviour was begun. During this and subsequent phases of the experiment the animals were maintained on a schedule by which they were fed M.R.C. Diet No. 41 (Bruce and Parkes, 1949) for 1 hour of each 24 hours. They were allowed an *ad libitum* supply of water in their home cages. Preliminary training followed the procedure suggested by Rabinovitch and Rosvald (1951) and took place in their modification of the original closed field test (Hebb and Williams, 1946). The apparatus consisted of an enclosed rectangular field with starting and food boxes at opposite ends of one diagonal. Movable wooden barriers could be placed at various positions in the field thus constructing the various spatial problems standardized by Hebb and Williams. Preliminary training took place in two phases. In the first, animals were put into the apparatus in pairs for one 15-minute period on each of five successive days, being allowed to wander through the field and to eat in the food box. In the second phase, each animal separately was released from the starting box on nine successive trials each day, a trial ending when the animal entered the food box. During both phases one of a set of five standardized practice problems was constructed in the field for each day's runs, the five problems being repeated in rotation until the preliminary training was completed. This training was concluded as soon as all animals had reached a criterion of

nine trials in 60 seconds or less on two successive days. Trials on the standardized set of experimental problems then began, each animal being given seven trials per day on one problem. Errors were recorded for entrance into "error zones" (Rabinovitch and Rosvald, 1951). This phase of the experiment was completed in twelve days, at which time data were available for each animal's reactions to twelve standardized problems.

In Phase 3 each animal was required to run a straightaway formed by barriers leading directly from the starting to the food box. The time taken for each of seven runs on three successive days provided data on speed of locomotion.

Experiment 2. This experiment was conducted in two replications designed to assess the effects of two levels of hypothermia on, first the retention of a previously practised habit, and, second, the serial problem-solving behaviour of the kind studied in the first experiment. The experimental design is shown in Table II. As in Experiment, 1, the various groups differed only in the treatment they received during Phase 3 when the deep body temperature of Group ES was reduced to $+1^{\circ}$ to 0° C. and that of Group EM to $+13.4^{\circ}$ to $+18.5^{\circ}$ C, Group C serving as a control. A total of forty animals was used, twenty in each replication. Six animals did not complete the full experiment, five as a result of the severe hypothermia. All animals were from the same stock as those used in the first experiment and were 62 days old at the start of Phase 1. Animals were assigned to the various groups in accordance with a litter-mate technique

TABLE II
EXPERIMENT 2. EXPERIMENTAL DESIGN

Group	Phase						
	1	2	3	4	5	6	7
ES	Pre-liminary Training	Maze Learning	Hypothermia: $+1^{\circ}$ to 0° C.	Maze Retention	Pre-liminary Training	Serial Problem Solving	Speed of Locomotion
EM	Pre-liminary Training	Maze Learning	Hypothermia: $+13.4^{\circ}$ to $+18.5^{\circ}$ C.	Maze Retention	Pre-liminary Training	Serial Problem Solving	Speed of Locomotion
C	Pre-liminary Training	Maze Learning	No Hypothermia	Maze Retention	Pre-liminary Training	Serial Problem Solving	Speed of Locomotion

The first part of the experiment was concerned with training in establishment of a maze habit, the retention of which could be tested after the experimental groups had been subjected to hypothermia. The Lashley No. III maze, adapted for use in water (Braun, Russell and Patton 1949), served as the problem to be learned. Preliminary training consisted of swimming a straightaway to escape from the water, each animal being given five trials per day for five days. Learning of the maze proper began on the next day. All animals were given an equal amount of training—one trial a day for eight days.

Two days later the experimental animals were cooled in accordance with the requirements of the experimental design. Thirteen days after this treatment tests for retention were started. The retention and savings methods were used, the animals being required for purposes of the latter method to swim the maze to a learning criterion of two successive days without error. As in the case of the original learning each animal received one trial per day.

After completing the test for retention each animal was put on a schedule of 1 hour feeding per 24 hours. Twelve days after the last animal had reached the learning criterion in the maze, training on the serial-problem-solving situation began. This was 45 days after the experimental animals had been subjected to hypothermia. The same procedure as in the first experiment was followed, yielding information about both problem solving and speed of locomotion.

In analysing the results significances of all differences among groups were tested by the Mann and Whitney (1949) method. "Significance" is defined in terms of the 5 per cent. level of confidence.

III

RESULTS

The two experiments were designed to study the effects of hypothermia on learning and on retention, and the results may be summarized under these two headings. However, there is certain general information which should be considered before these summaries are presented.

Speed of locomotion was measured in both experiments since it was possible that the experimental treatments might have affected locomotion differentially in the various groups and this, in turn, might have been reflected in the measures of learning. Comparisons of times taken to run the straightaway showed no significant differences among any of the groups in the two experiments. This suggests that any differences which might appear in the learning or memory scores of the various groups could not be accounted for in terms of inherent differences in speeds of locomotion.

The possibility that differences might have existed in levels of food-hunger drive among the various groups was also examined. The measure used was an indirect one—amount of food intake. In neither experiment were there any significant differences during the periods in which the various behaviour measures were being taken.

Effects on learning. The serial problem solving task provided a total error score as the main measure of learning. Such scores were available for all animals in both experiments. It must be remembered, however, that animals in the second experiment were tested for retention of previous practice on the Lashley III water maze before beginning training on the serial problem solving task.

TABLE III
SERIAL PROBLEM SOLVING—TOTAL ERROR SCORES

<i>Experiment</i>	<i>Group</i>	<i>N</i>	<i>Median</i>	<i>Range</i>
1	C	7	173.0	115-234
	ES	6	216.5	193-258 ¹
2	C	9	160.0	108-264
	EM	9	167.0	112-231
	ES	14	190.5	137-280
2	C + EM	18	164.5	108-264 ²
	ES	14	190.5	137-280 ²

^{1,2} These differences are significant in a one-tail test at the 0.03 and 0.05 levels of confidence respectively.

Table III presents the median error scores for the various groups in both experiments. There is a consistent trend towards greater errors for the groups which had been subjected to severe hypothermia. The difference between the two groups in Experiment 1 is significant ($p = 0.03$). Comparisons of the three groups in Experiment 2 provide further information about the relation between level of body cooling and learning. The difference between the control animals and those whose body temperature had been lowered to between $+13.4^{\circ}$ and 18.5° C. is far from significant ($p = 0.48$), whereas the differences between these two groups and the group which had been cooled to between $+1$ and 0° C. approached significance ($p = 0.07$ and 0.09 , respectively). When the error scores of the C and EM groups are combined

and then compared with the scores of the ES group the difference is significant ($p = 0.05$).

Effects on retention. The second experiment was designed to provide information regarding the effects of two levels of hypothermia on the retention of a relatively simple habit practised before body temperature was lowered. Although the three groups involved in each of the two replications were matched according to a litter-mate technique, it was considered desirable to compare their original practice scores in case differences at this point in the experiment may have affected the measures of retention. The median error scores for the C, EM, and ES groups were 42.5, 39.0 and 42.0, respectively. Mann and Whitney tests (1949) showed no significant differences among the groups.

TABLE IV
RETENTION FOLLOWING HYPOTHERMIA. EXPERIMENT 2

Group ¹	N	Errors on First Trial		Total Trials		Total Errors		Per cent. savings	
		Median	Range	Median	Range	Median	Range	Median	Range
C	10	1.0	0-17	2.0	0-12	3.0	0-59	90.5	34-100
EM	9	2.0	0-13	4.0	0-14	7.0	0-47	83.0	0-100
ES	15	2.0	0-8	7.0	0-18	8.0	0-82	88.0	19-100
C + EM	19	1.0	0-17	2.0	0-14	4.0	0-59	91.0	34-100
ES	15	2.0	0-8	7.0	0-18	8.0	0-82	88.0	19-100

¹None of the differences among these groups is significant.

Table IV summarizes the data on retention. The animal's error score on the first trial in the maze following Phase 3, the "hypothermia" phase, was taken as one measure of its retention. Comparisons of the differences between the various groups in both replications showed no significant differences, although there were slight tendencies for the control groups to make fewer errors than the groups previously subjected to hypothermia.

By continuing the training to a learning criterion of two successive errorless trials three additional, though interrelated, measures of retention were obtained: total trials and errors to reach the criterion and the "per cent. savings" expressed as the difference between the errors during original learning and the errors during re-training divided by the original error score. Of all possible comparisons among the various groups in the two replications none of the differences was statistically significant. However, examination of the various comparisons does indicate a consistent trend toward poorer retention by the groups whose deep-body temperature had previously been lowered than by the control groups.

IV

DISCUSSION

These results are of interest in four ways. First, they support the evidence reported by Ransmeier (1953) that severe hypothermia, which does not follow immediately after practice, has little or no effect on measures of memory. Although our results show a consistent trend toward poorer retention following hypothermia, the differences among experimental and control groups are very small and in no instance are they statistically significant. The small differences may be functions of the technique used to reduce deep body temperature rather than of the effects of



hypothermia *per se*. These results are difficult to explain if long-term memory is dependent upon the continuous activity of the brain since in our rats cerebral electrical activity was arrested for perhaps as long as $1\frac{1}{2}$ to 2 hours.

The second main point of interest concerns the effects of hypothermia on learning. Here, as in the case of retention, the results were consistent. The performance of rats whose body temperatures had been lowered to between $+1^{\circ}$ and 0° C. showed an impairment when compared with the performances of control animals, an impairment which proved to be statistically significant. On the other hand, no significant differences were found between the performances of control rats and those cooled to between $+13.4^{\circ}$ and $+18.5^{\circ}$ C. despite the very severe physiological changes induced in the latter. Thirteen to fifteen degrees Centigrade appears again as a "crucial point"; below this point such physiological processes as heart beat and respiration are arrested and performance on later learning tasks is impaired. Further examination of our results suggests that this impairment in learning may be only temporary, although additional research is necessary before this general statement can be made with confidence. Animals in Experiment 1 started training on the problem solving task 19 to 21 days after the "hypothermia" phase of the experimental design, while this interval for animals in Experiment 2 was 45 days. In terms of median performance scores, impairment in the first experiment amounted to 25 per cent. and impairment in the second to 16 per cent. This picture of a significant but temporary impairment in learning performance has been observed in studies of other forms of treatment involving drastic changes in physiological processes (e.g., Braun, Russell and Patton, 1949).

A third special feature is that our results are based upon observations of the behaviour of non-hibernating mammals. With such animals extreme hypothermia, such as that employed in the present study, results in complete arrest of circulation and respiration. In this respect they differ from mammals in a state of natural hibernation, and from hibernating or non-hibernating mammals artificially cooled to induce surgical anaesthesia, in all of which circulation and respiration continue. Our results have a different physiological significance from those of Ransmeier, who based his observations on the behaviour of hamsters cooled to temperatures which they would reach in normal hibernation and at which heart beats and breathing were slowed but not stopped. Within the limits of our experimental procedure we have failed to find evidence that arrest of vital processes as a result of hypothermia produces any serious effects on an animal's behaviour once it has been successfully reanimated.

A fourth matter of interest, particularly for the medical research worker, is the lack of evidence that hypothermia, even considerably more severe than that used in surgical procedures, has any serious, permanent effect on learning or retention. This point is, of course, limited to the types of behaviour we have studied and it would seem important to assess the possible effects of hypothermia on other types of behaviour and other species of animals.

ACKNOWLEDGEMENT •

We wish to express our thanks to the Medical Research Council for their support of this research.

REFERENCES

- ADOLPH, E. F. (1950). Oxygen consumptions of hypothermic rats and acclimatization to cold. *Amer. J. Physiol.*, **161**, 359-73.
 ANDJUS, R. K. (1951). Sur la possibilité de réanimer le rat adulte refroidi jusqu'à proximité du point de congélation. *C. R. Acad. Sci., Paris*, **232**, 1591-93.

- ANDJUS, R. K. (1955). Suspended animation in cooled, supercooled and frozen rats. *J. Physiol.*, **128**, 547-56.
- ANDJUS, R. K., and LOVELOCK, J. E. (1955). Reanimation of rats from body temperatures between 0° and 1° C. by microwave diathermy. *J. Physiol.*, **128**, 541-46.
- ANDJUS, R. K., and SMITH, A. U. (1954). Revival of hypothermic rats after arrest of circulation and respiration. *J. Physiol.*, **123**, 66-7 P.
- ANDJUS, R. K., and SMITH, A. U. (1955). Reanimation of adult rats from body temperatures between 0° and $+2^{\circ}$ C. *J. Physiol.*, **128**, 446-72.
- BRAUN, H. W., RUSSELL, R. W., and PATTON, R. A. (1949). Duration of decrements in learning and retention following electroshock convulsions in the white rat. *J. comp. physiol. Psychol.*, **42**, 87-106.
- BRUCE, H. M., and PARKES, A. S. (1949). Feeding and breeding of laboratory animals. IX. A complete cube diet for mice and rats. *J. Hyg., Camb.*, **47**, 202-8.
- BURTON, A. C., and EDHOLM, O. G. (1955). *Man in a Cold Environment*. London.
- CRISMON, J. M. (1944). Effect of hypothermia on the heart rate, the arterial pressure and the electrocardiogram of the rat. *Arch. intern. Med.*, **74**, 235-43.
- FAZEKAS, J. F., and HIMWICH, H. E. (1939). Effect of hypothermia on cerebral metabolism. *Proc. Soc. exp. Biol., N.Y.*, **42**, 537-38.
- GERARD, R. W. (1953). What is memory? *Sci. Amer.*, **189**, 118-26.
- HEBB, D. O., and WILLIAMS, K. (1946). A method of rating animal intelligence. *J. genet. Psychol.*, **34**, 59-65.
- HORSTEN, G. P. M. (1949). Influence of the body temperature on the E.E.G. *Acta brev. néerl. Physiol.*, **17**, 23-5.
- HUNTER, W. S. (1932). The effect of inactivity produced by cold upon learning and retention in the cockroach *Blatella germanica*. *J. genet. Psychol.*, **41**, 253-66.
- JENSEN, J. M., and PARKINS, W. M. (1954). Brain tolerance to differential hypothermia and circulatory occlusion. *Fed. Proc.*, **13**, 75.
- LEMAITRE, M. (1954). Der Einfluss der Temperatur auf das Elektrocorticogramm der Ratte. *Z. Biol.*, **106**, 426-35.
- MANN, H. B., and WHITNEY, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *Ann. math. Statist.*, **18**, 50-60.
- RABINOVITCH, M. S., and ROSVALD, H. E. (1951). A closed-field intelligence test for rats. *Canad. J. Psychol.*, **5**, 122-28.
- RANSMEIER, R. E. (1953). The effects of convulsion, hypoxia, hypothermia and anesthesia on retention in the hamster. Ph.D. Thesis, University of Chicago.

THE LIMITS OF THE "PSYCHOLOGICAL REFRACTORY PERIOD"

BY

R. DAVIS

*From the Institute of Experimental Psychology,
University of Oxford*

Previous evidence on the psychological refractory period is shown to be inadequate on the crucial issue of whether delays, which cannot be overcome by practice, occur in responding to the second of two signals, when the interval between them is less than 0.5 seconds. For simple key pressing responses it is shown that when the interval between signals is less than the reaction time to the first signal, delays in the reaction time to the second signal occur in a predictable manner. When the interval is greater than the first reaction time, however, no such delays are found. Possible reasons for the discrepancy between these and earlier findings are suggested.

INTRODUCTION

In many complex skills it is necessary for the operator to make repeated adjustments of his motor performance in accordance with changes in the stimulus situation. One of the most fundamental limitations of such performance is the rate at which responses can be made to signals from the environment.

The normal reaction time (of about 0.2 seconds for a visual stimulus) imposes one limitation, but claims have been made as a result of work by Craik (1948), Hick (1948), Vince (1948), and others that there is a further limitation on the rate of response; that, if a second signal follows a previous signal at an interval of 0.5 seconds or less, the beginning of the second response is delayed. This increase in the reaction time to the second stimulus, it is claimed, may be as much as 0.3 seconds. The term "psychological refractory period" has been given to the period following the presentation of the first stimulus during which a second stimulus would give rise to a delayed second reaction.

The refractory period was supposed to be of general occurrence in making any type of response to signals at short intervals. Thus it has been claimed that similar delays occur whether the subject moves the same or different body members in response to the consecutive stimuli (Hick, 1948; Hick & Bates, 1950). Of the variety of situations which have been covered by the assumption of a refractory period it has been applied to:—

- (a) "Ungraded" responses. Situations in which the subject had to press a key in response to light signals.
- (b) "Graded" responses. Situations in which the subject had to make a movement of definite extent (Vince, 1948 (a)). In this case the response to the second signal was a movement in the opposite direction to the first, back to the starting point.
- (c) The refractory period was further applied to a situation in which a target had to be tracked continuously. Misalignments between target and cursor lead to corrections being made. The frequency of such corrections has been reported to have an upper limit of approximately 2 per second and the hypothesis of the refractory period has been used to explain why this rate is not exceeded (cf. Craik, 1948; Welford, 1952).

Any delay which occurs must be a result of some characteristic of the organization of the central nervous system. It could hardly be determined either by transmission from the sense organ—the eye for a visual stimulus, which can respond at a much higher rate, or by the muscles or limbs where voluntary movements of the order of 10 per second are possible.

The hypothesis of some sort of central delay in dealing with sequences of signal had been suggested by Telford (1931) and elaborated by Craik (1948), but it is to Hick (1948) and Welford (1952) that we owe the development of the theory of the psychological refractory period in its most complete form as a physiological limit on the rate of responding.

A brief summary of the points of the theory as developed in Welford's article will be attempted in order to point the significance of the present investigation.

Delays in responding to a stimulus which closely follows another are due

"to the central processes concerned with two separate stimuli not being able to coexist, so that the data from a stimulus which arrives while the central mechanisms are dealing with data from a previous stimulus have to be 'held in store' until the mechanisms have been cleared." (Welford, 1952, p. 3.)

If we apply this to the central processes occurring during one reaction time only, then with two consecutive stimuli S_1 and S_2 each indicating a different movement M_1 and M_2 and separated by an interval I ($< RT_1$) we have the situation shown in Figure 1A.

RT_1 = Actual 1st reaction time

RT_2 = Predicted 2nd reaction time

RT_N = Mean normal reaction time

x = Amount of delay

Then $x = RT_1 - I$

$RT_2 = RT_N + x$

$\therefore RT_2 = RT_N + RT_1 - I$

The mean normal reaction time RT_N is the assumed mean reaction time to the second stimulus if the first stimulus had not in fact preceded it. Its value would be taken as the mean simple reaction time for the particular experimental situation used. Actual values of RT_2 should not differ from those predicted by the formula by more than the normal amount for this situation.

However, this formula will not explain delays when $I > RT_1$ which are claimed to occur for values of I up to 0.8 seconds. In order to explain these delays further assumptions are introduced. Thus

"The attention may be switched to that sensory field from which confirmation of the occurrence of the response is expected . . . we may suppose that the attention is reflexly deflected by the inevitable stimulation of kinaesthetic or other receptors by the response." (Hick, 1948, p. 43.)

Welford adds to this:

"If the stimuli fed back from a response are dealt with by the subject's central mechanisms in a similar way to stimuli given by the experimenter, we should expect that those fed back at the beginning of the response would 'capture' the central mechanisms for a brief period, and that only when this period was over could any further stimuli from the response or any new stimulus given by the experimenter be attended to." (Welford, 1952, p. 4.)

Thus the perception of stimuli resulting from the response would be intermittent and give rise to one or more periods of refractoriness.

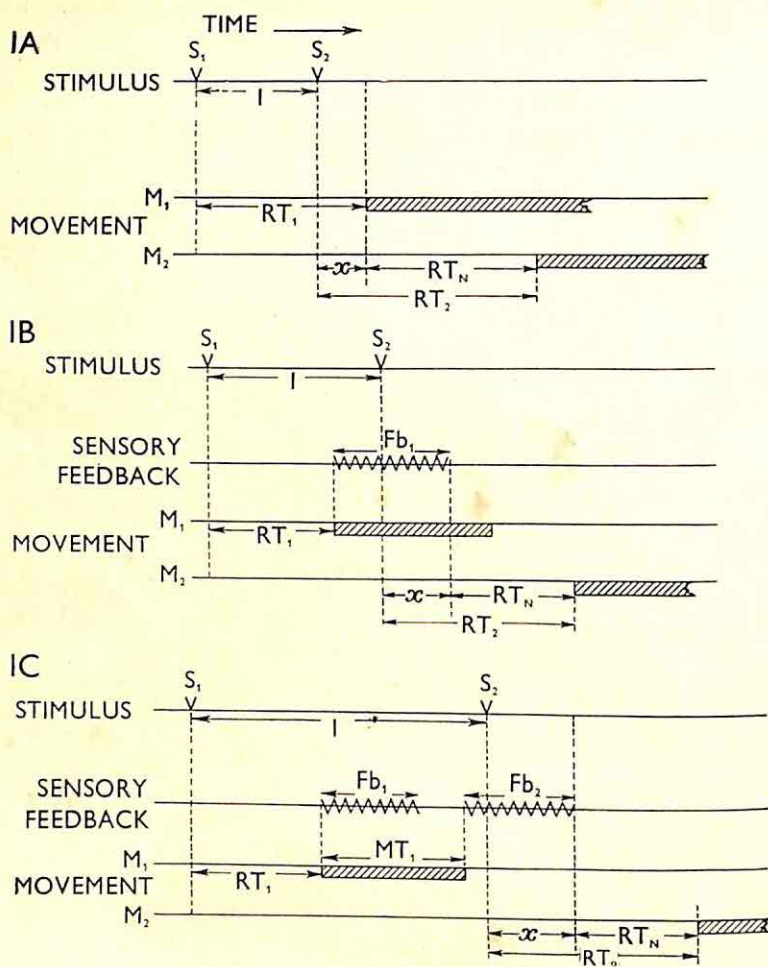
This situation is shown in Figures 1B and 1C.

Fb_1 = first period of sensory feedback MT_1 = first movement time.

Fb_2 = second period of sensory feedback.

Bureau Ednl. Psy. Research
DAVID H. TRAINING COLLEGE
Dated
Acce. No

FIGURE 1



Showing the derivation of predicted values of RT_2 for different values of interval between the first and second stimuli.

- (IA) When $RT_1 > I$
 $RT_2 = RT_N + RT_1 - I$
- (IB) When $RT_1 + Fb_1 > I > RT_1$
 $RT_2 = RT_N + RT_1 + Fb_1 - I$
- (IC) When $RT_1 + MT_1 + Fb_2 > I > RT_1 + MT_1$
 $RT_2 = RT_N + RT_1 + MT_1 + Fb_2 - I$

I is Interval between stimuli S_1 and S_2

RT_1 = Reaction time to 1st stimulus S_1

RT_2 = Predicted reaction time to 2nd Stimulus S_2

RT_N = Mean normal reaction time (= Mean RT_1)

MT_1 = Duration of 1st movement M_1

Fb_1 = duration of 1st period of sensory feed back

Fb_2 = " " 2nd " " " " "

x = Amount of delay.

This gives rise to the following values of RT_2 for different values of I :

Figure 1B: When S_2 occurs at a point such that $RT_1 + Fb_1 > I > RT_1$

$$\text{Then } I + x = RT_1 + Fb_1$$

$$\text{And } RT_2 = RT_N + x$$

$$\therefore RT_2 = RT_N + RT_1 + Fb_1 - I$$

Figure 1C: When S_2 occurs at a point such that $RT_1 + MT_1 + Fb_2 > I > RT_1 + MT_1$

$$\text{Then } I + x = RT_1 + MT_1 + Fb_2$$

$$RT_2 = RT_N + x$$

$$\therefore RT_2 = RT_N + RT_1 + MT_1 + Fb_2 - I.$$

With regard to the values of Fb Welford assumes that it is the result of the inevitable kinaesthetic stimulation and takes values of approximately 0.15 seconds as determined by kinaesthetic reaction time experiments (Vince, 1948 (b)). Such feedback is supposed to occur at "high points" in the movement and is not necessarily confined to the beginning and end of the movement, although in a straightforward translative movement Welford gives these as the most likely places.

This provides a precise theory in an admirable form for testing experimentally. Taken as it is, the order in which the external and internal stimuli occupy the central mechanisms will be determined solely by the time relations between them, i.e. the sequence in which they arrive, and the duration for which they occupy the central mechanisms will result in observed delays. At times, however, Welford seems to propose a much weaker form of the theory wherein the order in which the stimuli occupy the central mechanisms is not determined solely by the sequence of their arrival. The central mechanisms are given the power of "paying, or not paying, attention" to certain stimuli (Welford, 1952, pp. 16-17). This obviously would raise further problems concerning the selective behaviour of the central mechanisms and robs the theory of much of its predictive power.

However, an experimental test of the theory in either its precise or its weaker form may be designed as follows:—

If the feedback stimuli, whatever they are, are occupying the central mechanisms for a certain time, then, if the experimenter gives a signal to respond during this time, it is of no advantage giving it at the beginning of this period rather than at the end, since it is not dealt with until the end of this period in any case. Hence shortening the interval by this amount will not produce any shortening in the second reaction time. Thus if an experiment is performed in which the intervals between the experimentally produced stimuli are varied, and the steps between the intervals are sufficiently small compared with the suggested value of feedback times we should discover—

(1) Whether there is any delay in the second reaction for values of the interval (a) less than RT_1 , (b) between RT_1 and 0.5 seconds.

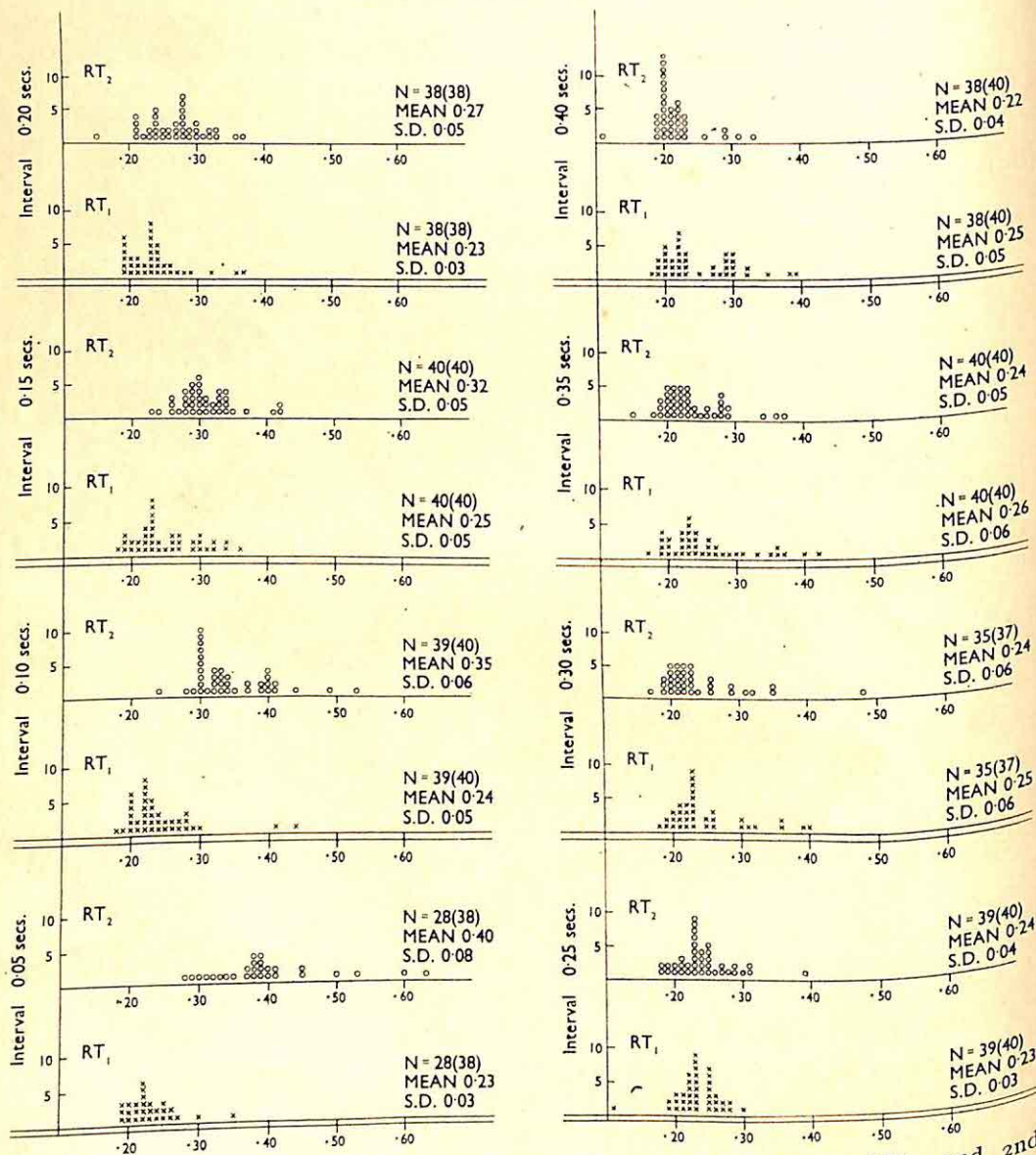
(2) By observing how much the interval can be lengthened without producing further delay we can obtain a value for the time sensory feedback is occupying the central mechanisms.

Although a large number of experiments have been quoted in support of a refractory period no precise analysis of this sort seems to have been attempted. The majority of studies have merely reported a delay of some sort at short intervals. Thus the early experiments of Telford used only four values of interval, none less than $\frac{1}{2}$ second, viz.: $\frac{1}{2}$, 1, 2, 4 seconds. Hick's results in a key-pressing task (1948) give only the mean second reaction time (0.389 seconds) of 14 cases for all intervals grouped together from 0.171 to 0.513 seconds. Vince's results (1948a) it is true are given for

values of interval 0.05, 0.10, 0.20 . . . 0.7 seconds, but this is a situation in which the two movements are performed by the same limb in opposite directions, and as movement time plus reaction time totals approximately 0.5 seconds, as Vince herself points out (1948a, p. 156) this may be all that is required to account for the time taken to return, without introducing a refractory period.

It therefore seemed necessary to perform an experiment of the type outlined to provide the basic data in a simple key-pressing situation where the performance of the first movement was unlikely to interfere with the second.

FIGURE 2 (A) SUBJECT I



Distribution of 1st reaction times, RT₁, and 2nd Reaction times in seconds to the nearest 1/100 second along the horizontal. Arithmetical means and standard deviations are also given. N denotes the number of excluding failures in the ink-

EXPERIMENTAL PROCEDURE

Apparatus.

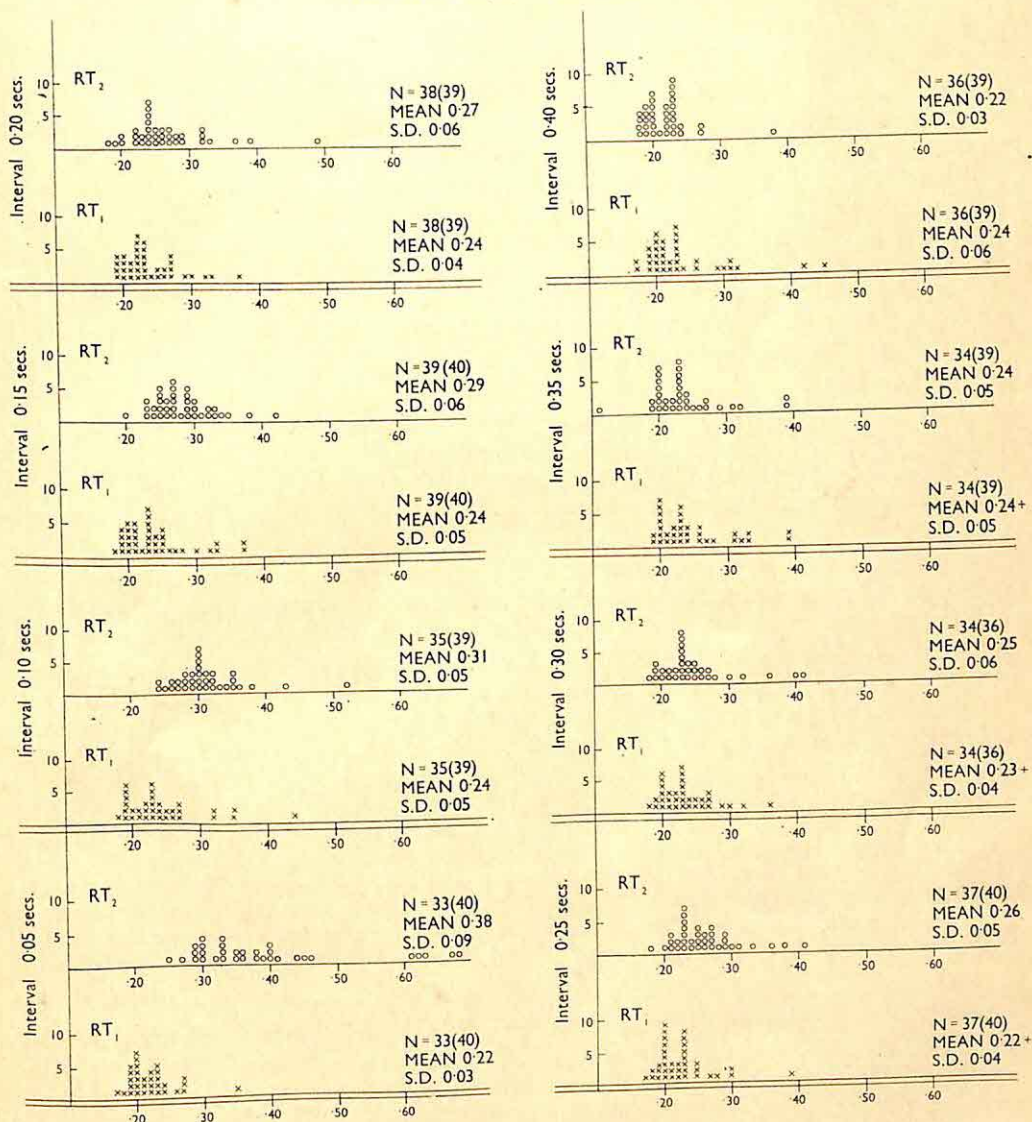
The subject viewed an illuminated moving band passing behind a screen with a narrow slit and his task was to press a key whenever a spot was exposed in the slit. Two keys were provided for the subject and two conditions were used:—

1. In which the two keys were on opposite sides of the display and were pressed by the same fingers of *different* hands.

1st reaction—1st finger left hand;

2nd reaction—1st finger right hand.

FIGURE 2 (B) SUBJECT 2



reaction times, RT₂, for intervals 0.05 — 0.40 seconds.

axis. The number of responses falling in each category along the vertical axis.

pairs of responses made at each interval. The total possible number of pairs of responses recording, is shown in brackets.

2. In which the two keys were adjacent and were pressed by different fingers of the same hand.

1st reaction—1st finger right hand;

2nd reaction—2nd/3rd fingers right hand.

When a key was depressed an electromagnetic pen recorded on the paper band and reaction times were measured to the nearest $1/100$ second as distances along the band from the stimulus spot to the pen's marking. The reaction times were measured from the moment the stimulus first appeared at the top of the slit, not after the spot had crossed the slit as in some of Vince's experiments (1948a).

Numerical Specifications of Apparatus.

Horizontal side of slit. 2 cms.

Vertical side of slit 2.5 mms.

Approx. distance of subject's eye from slit

20–30 cms.

Size of stimulus spot 8 mms. (horizontally) $\times$ 2 mms. (vertically).

Speed of paper. 4.92 ± 0.03 cms./sec.

Force necessary to operate key Approx. 80 gms. weight.

The paper was illuminated by a 6 v. bulb attached behind the screen immediately above the slit.

METHOD

Main Experiment.

When the paper band was set in motion the subject first received a warning signal in the form of a horizontal line drawn across the band.

At a randomly varied time between 1.5 seconds and 2.5 seconds after this the first stimulus appeared through the slit. As soon as the subject saw this he had to press the first key.

At a randomly varied time between 0.05 and 1.00 second after the first stimulus a second stimulus might appear. When this happened the subject had to press the second key.

After this another warning signal would appear and the procedure was repeated.

Sixty-four pairs of stimuli were presented on one paper band, and distributed randomly among these pairs were 12 cases in which there was a single stimulus only. These served to check whether the first reaction was being performed normally and whether there was a tendency to anticipate the second reaction.

Each subject was given 128 pairs of stimuli per session (i.e. 2 paper bands), with a short rest after every 32 pairs. Five sessions were held in all with intervals of about two days between them.

Control Experiment.

In separate control experiments the simple reaction times for different hands, or for different fingers of the same hand, were determined according to the condition under which the subject performed the main experiment, except that the subject knew that following the warning signal there would be only one stimulus to which he had to respond.

RESULTS

Results are given for two subjects, both of whom had previously performed several hundred responses in situations similar to those of the experiment—an essential precaution when one is looking for absolute limitations on performance. For *subject 1* condition 1 was used—two keys pressed by the same fingers of different hands.

For *subject 2* condition 2 was used—two keys pressed by different fingers of the same hand.

In Figure 2 are shown the distributions of first and second reaction times for intervals from 0.05 to 0.40 seconds. On these distributions every complete pair of

responses has been shown, no matter how atypical. There was a possible total of 40 pairs of responses at each interval. Failure to respond accounts for about three-quarters of the missing pairs, most of these failures occurring at the 0.05 seconds interval. Failure in the ink recording accounts for the rest.

From this figure *five* points might be noticed.

1. When the interval between the first and second stimuli is less than 0.25 seconds the distributions of the second reaction times show a delay compared with the corresponding distribution of the first.

2. When the interval is greater than 0.25 seconds no such shift of the distribution is evident—although occasional long values occur in both first and second reaction times.

TABLE I

		MEAN	S.D.	N	For single reactions in main experiments		
					No. correctly performed	Possible No.	Per cent. correctly performed
SUBJECT 1	1st finger, left hand	0.23	0.03	110			
	1st finger, right hand	0.24	0.05	124			
	Single reactions in main experiment	0.25	0.04		100	117	86
SUBJECT 2	1st finger, right hand	0.24	0.04	121			
	2nd/3rd finger right hand ..	0.22	0.02 +	131			
	Single reactions in main experiment	0.23	0.04		104	113	92

Arithmetical means and standard deviations for simple reaction times in the control experiment. Also included are the means and standard deviations for the simple reactions which were interspersed through the main experiment, with the number correctly performed as simple reactions out the possible total.

3. As the interval between stimuli increases from 0.05 second by 0.05 second steps, the shift in the distribution of second reaction times is reduced by approximately 0.05 steps.

4. Subjects performing under the two different conditions (same or different hands) give very similar results.

5. Apart from the abnormally large scatter (S.D. = 0.08 for subject 1 and 0.09 for subject 2) for the second reaction times at interval 0.05 seconds, the type of distribution is similar for first and second reactions at all intervals.

At 0.05 seconds there may be a combination of disturbing factors contributing to the large scatter—as well as to the smaller number of completed pairs of responses. 0.05 seconds is about the time between successive fixations of the eye, and the vertical

width of the slit was such that it took the stimulus 0.05 seconds to cross it. If the eye tended to follow the stimulus spot down and then jump back there was frequently some uncertainty about whether one or two spots had been seen. This, however, was the only interval at which either subject expressed any doubt about the clarity of the stimulus.

In comparing the distribution of first and second reaction times the assumption is made that the first reaction times are normal—i.e. that they are similar to those obtaining for a simple reaction under these conditions.

The results of the control experiment serve to check this. In Table I are given means and standard deviations for the subjects performing a simple reaction with the limb indicated.

All distributions of first reaction times in the main experiment show very similar values, giving an approximate mean first reaction time $RT_N = 0.23$ seconds.

This may also be compared with the values of the reaction times to the single stimuli which were interspersed among the pairs in the main experiment.

Also shown in Table I are the means, S.D.'s and number of these reactions correctly performed, i.e. without any second response following the first. These again give an approximate mean value of 0.23 seconds.

It has been pointed out that highly practised subjects were used so that no improvement in reaction times was expected during the experiment. This was confirmed by comparing for each subject at each interval between 0.05 seconds and 0.25 seconds the mean first reaction times during the first two sessions with the mean first reaction times during the last two sessions (five sessions in all). Mean second reaction times were compared in a similar manner.

In no case was there any evidence for an effect attributable to practice during the experiment.

In order to explain certain apparent exceptions to the refractory-period theory a process called "grouping" has been suggested. This amounts to receiving two consecutive stimuli as a joint stimulus and performing a single complex reaction to them—thus reducing the apparent second reaction time.

This might be done in the experimental situation by delaying the initiation of the first response until the second stimulus had arrived and then making a double response. It seems unlikely that this is in fact occurring as the means and distributions of the first reaction times are very similar to those for simple reaction times with the same apparatus.

However, it is possible that second reaction times shorter than the mean value might occur in association with first reaction times longer than the mean value, so as a further precaution the product-moment correlation coefficients for the pairs of first and second reaction times were calculated. If there is such an association it should be revealed by a negative correlation between first and second reaction times.

Correlation coefficients are shown in Table II.

Thus except for the one value for Subject 2 at interval 0.10 seconds the correlations are *positive*, short second reaction times are associated with short first reaction times rather than the reverse, which one should expect if a grouping effect occurred.

The hypothesis we proposed to test was that when

$$I < RT_1$$

$$RT_2 = RT_N + RT_1 - I$$

and when $I > RT_1$

$$RT_2 = RT_N, \text{ i.e. no delay.}$$

This gives a set of predicted values for RT_2 at each interval.

TABLE II

	Interval (seconds)				
	0.05	0.10	0.15	0.20	0.25
SUBJECT 1:					
$n = N - 2$	26	35	38	35	36
r	+ 0.43	+ 0.66	+ 0.32	+ 0.26	+ 0.92
Whether significant at $P = 0.05$..	Yes	Yes	Yes	No	Yes
SUBJECT 2:					
$n = N - 2$	31	32	37	36	35
r	+ 0.59	- 0.64	+ 0.85	+ 0.21	+ 0.50
Whether significant at $P = 0.05$..	Yes	Yes	Yes	No	Yes

Product moment correlation coefficients, r , for the values of first and second reaction times in each pair of responses for intervals 0.05-0.25 seconds. Whether the values of r are significant at $P = 0.05$ is also shown.

N denotes number of pairs of responses at each interval.

In this and the subsequent table and figure any pair of responses showing highly atypical values has been excluded from the calculation.

The criteria for exclusion were:

- Any pair in which either reaction time was equal to or less than 0.15 seconds.
- Any pair in which the first reaction time was equal to or greater than 0.41 seconds.

This in fact only excludes 12 pairs out of the whole range of intervals for Subject 1 and 9 pairs for Subject 2.

In order to give a precise test of the hypothesis the formula has been applied to each pair of first and second reaction times for the intervals 0.05 to 0.25 seconds and a χ^2 test of significance for the difference between observed and predicted values has been performed.

The results of this are shown in Table III. Actual means are shown for comparison (see footnote referring to Table II).

Thus the only two cases which deviate from the predicted values are for Subject 2 at the shortest intervals. It is interesting to note, however, that the deviations result from the second reaction time being *shorter* than the predicted values. It was at interval 0.10 for Subject 2 that the only case of short second reaction times correlating positively with long first reaction times was found, so that it seems likely that some "grouping" effect may be taking place at this interval.

The divergences, although significant are not large, however.

Finally a graph has been plotted for the whole range of intervals showing the Means of the differences $RT_2 - RT_1$, and whether they are significant according to the "t" test.

This is shown in Figure 3.

The rapid and regular decrease of the difference between first and second reaction times for intervals less than 0.25 seconds may be noted and two further points deserve comment.

(1) Although for Subject 2 at an interval of 0.25 seconds the difference $RT_2 - RT_1$ is significantly positive, the values at this interval are still in agreement with the hypothesis of no delay when $I > RT_1$ —as is shown in Table III.

TABLE III

	Interval (seconds)				
	0.05	0.10	0.15	0.20	0.25
Predicted $RT_2 = RT_1 + RT_N - I$ ($I < RT_1$)	$RT_1 + 0.18$	$RT_1 + 0.13$	$RT_1 + 0.08$	*	*
SUBJECT 1:					
Actual Mean RT_2	0.40 (0.38)	0.34	0.32	0.27	0.24
χ^2 for differences between predicted RT_2 and corresponding actual RT_2	58.1 (32.6)	31.2	14.5	11.6	15.6
Number of pairs	28 (26)	35	40	37	38
Whether differences are significant at $P = 0.05$	Yes (No)	No	No	No	No
SUBJECT 2:					
Actual Mean RT_2	0.38	0.31	0.29	0.27	0.26
χ^2 for differences between predicted RT_2 and corresponding actual RT_2	56.5	62.6	31.6	47.2	36.1
Number of pairs	33	34	39	38	37
Whether differences are significant at $P = 0.05$	Yes	Yes	No	No	No

Test of the hypothesis that when $I < RT_1$, $RT_2 = RT_N + RT_1 - I$ and when $I > RT_1$, $RT_2 = RT_N = \text{Mean } RT_1$ for intervals 0.05 seconds - 0.25 seconds. Values of χ^2 for the differences between observed and predicted values are shown and whether these differences are significant at $P = 0.05$.

* At intervals 0.20 and 0.25 some of the values of RT_1 are longer, some shorter than the interval. The appropriate formula is applied in each case. As Subject 1's results included two atypically long values of RT_2 , viz.: 0.60 seconds and 0.64 seconds at interval 0.05 seconds, also shown in brackets in this column are the corresponding values when these two responses are excluded.

The significant difference in Figure 2 is accounted for by the fact that at the interval 0.25 seconds several of the first reaction times are in fact longer than the interval. For these values a delay in the second reaction time is still predicted and this results in the significant (but small) difference observed.

(2) After RT_2 has decreased to the approximate value of RT_1 at $I = RT_1$ it remains at about this value for increasing intervals until towards the middle of the range it tends to get slightly shorter than the first reaction time.

This is an effect of the type noted by Mowrer (1940) and will be considered further in the discussion.

DISCUSSION

That there is a progressively increasing delay in the second reaction time as the interval between stimuli is reduced to less than the first reaction time would seem to be clearly established, and quantitative considerations support the formula suggested fairly well. That this delay disappears as soon as the interval between the stimuli is greater than the first reaction time is also clear.

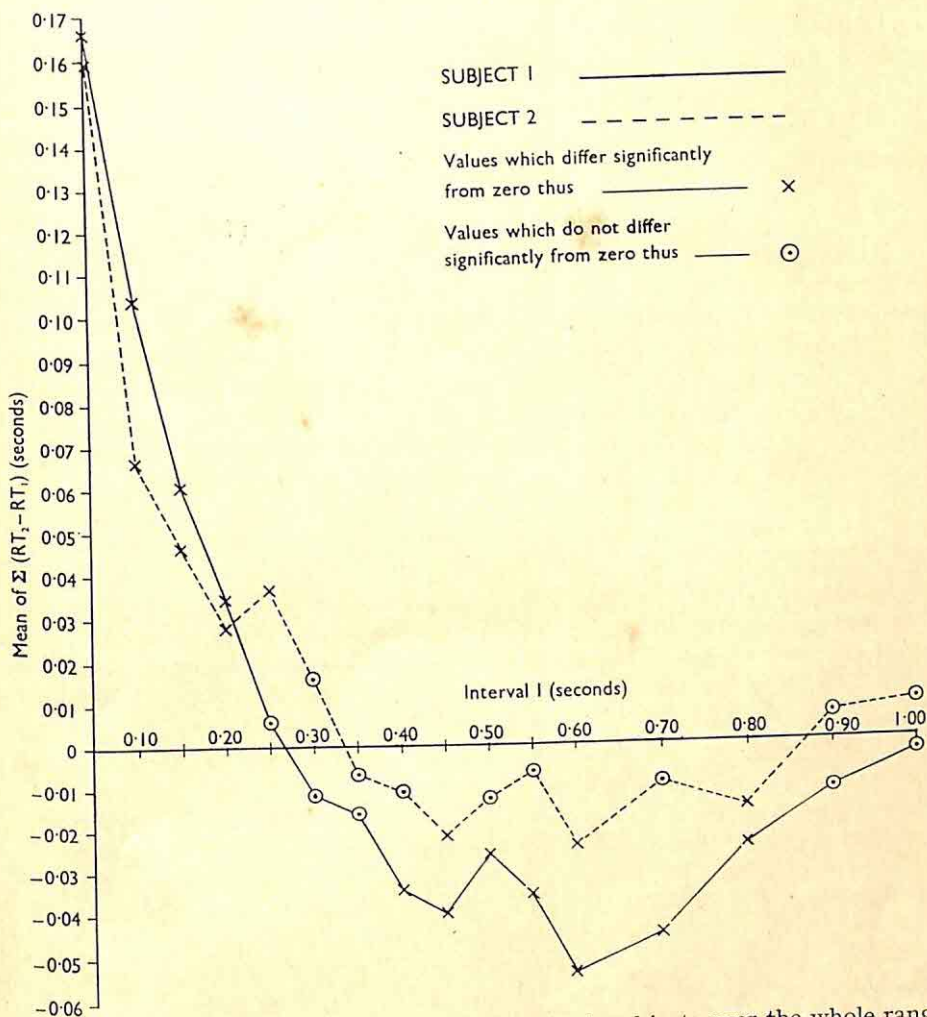
"Expectancy" Theories.

In view of the fact that some writers, most recently Elithorn and Lawrence (1955), have attempted to account for refractory-period observations in terms of the subject's

"expectancy" of responding at different intervals, it may be appropriate here to consider the significance of the slight shortening of the second reaction times with respect to the first in the interval range 0.40 to 0.80 seconds.

Subject 1 shows this most clearly, with RT_2 dropping to a minimum of 0.60 seconds and increasing again until it no longer differs from RT_1 at intervals 0.90 and 1 second.

FIGURE 3



Graph showing the mean of $\Sigma(RT_2 - RT_1)$ for both subjects over the whole range of intervals, and whether these values differ significantly from zero at $P = 0.05$ according to the "t" test.

Mowrer (1940) noticed that after subjects were trained to expect a stimulus at a certain interval from the previous one, then if the stimulus was given at some other interval the more this differed from the expected interval the greater was the reaction time.

There was no fixed optimal time interval which would result in a minimum reaction time to the second stimulus. The minimum was given at that interval which the subject had been trained to expect.

This would not give a prediction for a situation in which all intervals within a certain range were being reacted to equally frequently as in the present experiment. However, an expectancy theory might be extended to cover this situation as follows:—

1. The subject reacts most quickly at the interval to which he has been trained. Progressive deviations from this interval result in increasing delay.

2. Where several intervals are used in the training situation the subject reacts most quickly.

(a) at the most frequently presented interval;

(b) at the interval in the middle of the range.

There are obviously other alternatives.

None of these cases, however, deal with any absolute limit on the rate of response. The value of the reaction time depends on the interval or range of intervals used in training.

Elithorn and Lawrence (1955) have attempted to account for all delays occurring at short intervals in terms of expectancy.

It is difficult to compare their results with those in other refractory-period experiments as the first reaction was a disjunctive (two-choice) reaction with a fixed fore period instead of the more usual simple reaction with a variable fore period. Two of the three subjects also appear to be relatively unpractised; delays which have interested writers on the refractory period have been those persisting in the face of considerable practice. If the subjects are unpractised, as revealed by improvement during the experiment, this will tend to obscure the effects of an absolute limitation on performance.

In so far as Elithorn and Lawrence claim a maximum second reaction time at an interval between 50 and 250 milliseconds rather than at the shortest interval used (20 milliseconds) their results appear incompatible with both the above experiments and those of Hick and Vince—unless “grouping” of responses was occurring.

With regard to the theoretical interpretation, Elithorn and Lawrence use the cumulative normal probability curve over the range of intervals to predict the rise of expectancy—and consequently the reduction in the second reaction time. Thus, if, at any time after the first stimulus, the second stimulus has not occurred, then the probability that it will occur in the next moment of time will increase as the interval gets greater. This, however, is not the curve which fits Mowrer's results, nor will it apply to the intervals 0.40 to 0.80 seconds in the present experiment. In both these cases the expectancy follows the normal curve with maximum expectancy and minimum second reaction time at the centre of the range. The fact that Elithorn and Lawrence also obtain a minimum RT_2 at intervals 600–750 milliseconds rather than at their longest interval (950 milliseconds) would tend to support the normal curve of expectancy for their results, rather than their own view.

The results of the present writer's own experiments, however, are hardly amenable to an “expectancy” treatment over the whole range; they indicate a delay which cannot be overcome by well practised subjects fully aware of what is expected of them, when the interval between stimuli is less than the first reaction time. Once this value of the interval is exceeded other variations in reaction times may be a result of such factors as changing expectancy.

General Conclusions.

It has been established that discontinuity in the functioning of the operator exists, but not for so long a period as had been supposed. As the present results differ from those of earlier writers some suggestions might be made to account for the discrepancy.

It has already been pointed out that some of the early data were not really adequate evidence for the theory.

Furthermore previous writers on the refractory period have grouped together a large variety of response situations in their discussions. Thus key pressing tasks, situations where the subject has to make a graded movement of definite and known extent, and situations where the subject has to follow an unpredictable target have all been considered alike from the point of view of the delays in responding that might be expected.

Clearly, however, the control necessary for the execution of the movement may be very different in these cases.

For key pressing the amount of force applied to the key is relatively unimportant. As long as it exceeds a critical value the key will open. The "inevitable kinaesthetic stimulation" as a result of the movement may be present, but its value is unimportant for controlling the movement.

When a subject has to make a graded movement of definite extent from one point to another the accuracy of his movement will depend on the force he applies. In this case kinaesthetic feedback stimuli may be necessary to exercise this control.

When a subject has to follow an unpredictable target there are further complicating factors. After each movement of the subject the position of the cursor with respect to the target must be reassessed, usually visually, before the next move can be made. If, as has been claimed, corrections cannot be made more rapidly than about two per second, it might be this feature of the tracking task rather than occupation with kinaesthetic feedback stimulation which results in the limitation.

It might be argued that with the more complicated situations arising in various skills the delays may depend on the relative amount of information required to perform the task adequately.

However, experiments are in progress to investigate these situations further and results so far obtained would seem to indicate that in one further case at least, that is for graded movements of fixed extent, there is again no refractory period greater than the duration of one reaction time.

ACKNOWLEDGEMENTS

The writer wishes especially to thank Dr. H. Kay for his help and criticism at all stages of the work, as well as other members of the Oxford Institute of Experimental Psychology, and Professor G. Humphrey for affording the experimental facilities.

The work was undertaken as part of a project under a grant from the Conditional Aid Fund by the M.R.C. and D.S.I.R. Joint Committee on Individual Efficiency in Industry.

REFERENCES

- CRAIK, K. J. W. (1947). Theory of the human operator in control systems. I. The operator as an engineering system. *Brit. J. Psychol.*, **38**, 56-61.
- CRAIK, K. J. W. (1948). Theory of the human operator in control systems. II. Man as an element in a control system. *Brit. J. Psychol.*, **38**, 142-148.
- ELITHORN, A. and LAWRENCE, C. (1955). Central inhibition. Some refractory observations. *Quart. J. exp. Psychol.*, **7**, 116-127.
- HICK, W. E. (1948). Discontinuous functioning of the human operator in pursuit tasks. *Quart. J. exp. Psychol.*, **1**, 36-57.
- HICK, W. E. and BATES, J. A. V. (1950). The Human Operator of Control Mechanisms. Ministry of Supply Permanent Records of Research and Development. *Monograph* No. 17,204.
- MOWRER, O. H. (1940). Preparatory Set—some methods of measurement. *Psychol. Rev.* *Monog.* No. 52.

- TELFORD, C. N. (1931). Refractory phase of voluntary and associative responses. *J. exp. Psychol.*, **14**, 1-35.
- VINCE, M. A. (1948a). The intermittency of control movements and the psychological refractory period. *Brit. J. Psychol.*, **38**, 149-157.
- VINCE, M. A. (1948b). Corrective movements in a pursuit task. *Quart. J. exp. Psychol.*, **1**, 85-103.
- WELFORD, A. T. (1952). The "psychological refractory period," and the timing of high-speed performance—a review and a theory. *Brit. J. Psychol.*, **43**, 2-19.

Central Inhibition: Some Refractory Observations

TWO COMMENTS ON THE PAPER BY ALICK ELITHORN
AND CATHERINE LAWRENCE

(*Quart. J. exp. Psychol.*, 1955, 7, 116-27)

1.

By R. DAVIS

From the Institute of Experimental Psychology, University of Oxford

"The arrival of a single stimulus raises the probability of its own response to the region of 1.0, but at the same time it causes the probability of the alternative situation to fall abruptly. Since this latter situation will arise sometime the probability that it will occur in the next time interval if it has not already occurred will increase with the passage of time until towards the end of the cycle this probability approaches unity." (Elithorn & Lawrence, 1955, p. 125.)

The first probability here referred to is the probability of one or other of the stimulus-response patterns occurring first.

An entirely different probability is the probability of when in the time following the first, the second will occur. It does not follow that this probability is reduced to a relatively low value as a result of the first stimulus occurring.

There is no causal connection at all between the two probability measures. It is therefore a mistake to refer to some single probability level, changes in which are supposed to be reflected in changing excitability of the nervous system.

In seeking to explain the changes in the levels of probability by processes within the organism, if any one probability measure is taken, such as the changing probability of the second stimulus occurring as the time interval increases, then it might be possible to identify a state or process in the nervous system with the changing probability level.

If, however, a different set of probabilities are measured, it is absurd to expect that some process in the system we are examining will reflect the change in probability value, when this change is a result of changing the circumstances about which the probability statement is made.

Thus to talk of "the process of inhibition or whatever physiological change accompanies sudden alterations in probability level" (and which takes a finite time) involves the confusion of invoking a process to account for a supposed change in the organism when the change in fact lies in the choice of what is to be measured.

The assumption that the state of preparedness of the subject for the second stimulus (and hence reduction of reaction time) keeps increasing with the passage of time after the first stimulus has occurred seems in conflict with previously established evidence, e.g. Mowrer (1940) and also with Elithorn and Lawrence's own results. The minimum second reaction times do not in fact occur at the longest interval. This discrepancy has already been discussed more fully above.

REFERENCE

- MOWRER, O. H. (1940). Preparatory Set. Some Methods of Measurement. *Psychol. Rev. Monogr.* No. 52.

2.

By W. E. HICK and A. T. WELFORD

From the Psychological Laboratory, University of Cambridge

The authors deserve credit for confirming that psychological refractoriness, in the sense of the prolongation of the reaction time to the second of a pair of successive stimuli, does not depend on motor incompatibility between the responses. On the other hand, both their theorizing and their treatment of previous work are open to serious criticism.

(i) The authors attribute to Hick (1948), Vince (1948) and Welford (1952) a "theory which holds . . . that the integrative nervous activity of the organism is steplike and seriatim rather than diffuse and continuous." We cannot answer for Vince, but for ourselves we do not accept this. To import such terms as "integrative nervous activity" and "diffuse," which mean nothing in this context, is unwarrantable. It is the *decision process* that is claimed to be discrete and discontinuous.

(ii) The authors say: "Welford's theory implies that the delay . . . will be longest when the pair of stimuli are closest together. With Vince's data this is the case, but it is not so with the results of the present experiment." The truth is that Vince (*loc. cit.* and 1950) observed this seeming anomaly (she was probably the first to do so) and mentioned it perfectly clearly and explicitly. Moreover, Welford (*loc. cit.*) put forward three possible ways in which it might be brought to conform to the general terms of his theory.

(iii) The authors state that the inverse relation between the probability of a stimulus and the reaction time to it "was originally demonstrated by Merkel." If they will refer again to Hick (1952) they will see that the evidence is that Merkel had no idea that he was demonstrating any such thing. Moreover, as he did not say whether he presented the stimuli in any approximation to random order, it is not even certain that he can be credited with an unwitting demonstration.

(iv) The possibility of explaining refractoriness in terms of expectancy was discussed by Hick (1948), who pointed out that it would require an *ad hoc* hypothesis, for which he made certain tentative suggestions. Welford (1952) developed them in more precise form and showed that they could account for nearly all the data then available. It is true that he did not employ the term "expectancy," but his principle of non-overlap of central organizing times is an obvious consequence of the proportionality between such times and the corresponding amount of information to be organized (*see, e.g.* Hick (1952) and Hyman (1953)). For example, if two statistically independent stimuli have to be dealt with, the total organizing time will be the sum of their individual times, which will therefore appear to be incapable of overlapping.

The most important *ad hoc* hypothesis in Welford's theory is that, apart from some special cases, organizing time may be appropriated by the confirmatory feedback from the response. Since this feedback is essential to learning the skill, it would be gratuitous to suppose it is not also needed to maintain it, except perhaps for short periods and when accuracy can be sacrificed. In short, the feedback hypothesis is not strictly *ad hoc*, since it has some collateral support. (Further support comes from the work of Leonard discussed in Welford's paper.)

(v) Notwithstanding all this, Elithorn and Lawrence proceed, without acknowledgement, to state an expectancy theory which evades the need for an *ad hoc* hypothesis by introducing a logical fallacy. They say: "The arrival of a single stimulus raises the probability of its own response to the region of 1.0, but at the same time it causes the probability of the alternative situation to fall abruptly." It is surely obvious that the occurrence of the first stimulus raises the probability of the "alternative situation" to unity. The temporal uncertainty remains the same, of course. One has only to write down a list of the possible "messages" (right-left at all intervals plus left-right at all intervals) to see that the first stimulus reduces the prior uncertainty by one "bit." The fact that subjects cannot, or do not, use more than a small part of this information is what demands explanation, because it is what manifests itself as refractoriness.

(vi) The secondary problem of the reduced refractoriness at very short inter-stimulus intervals has been tentatively attributed by Vince and Welford to "grouping." To make this a little more precise, suppose the subject expects, in addition to the message-categories RL and LR, a third—namely R and L "together." This third category includes all intervals less than the least at which the *order* of two successive events is reliably distinguishable; it is not sharply defined, but is well known to be in the region of 50 to 100 ms. Assuming that it is sharply defined (which cannot make a serious difference on the average) and that the subject never confuses the three categories, misses a response, or responds prematurely, the information transfer will be $\log(3 + 1)$ approximately (Hick, 1952). Hence the reaction time for the (double) response to the third category should be about that for a normal three-choice situation. As far as can be judged, this is of the right order of magnitude. More cannot be said without a detailed analysis and, probably, a specially-designed experiment.

REFERENCES

- HICK, W. E. (1948). Discontinuous functioning of the human operator in pursuit tasks. *Quart. J. exp. Psychol.*, **1**, 36-57.
- HICK, W. E. (1952). On the rate of gain of information. *Quart. J. exp. Psychol.*, **4**, 11-26.
- HYMAN, R. (1953). Stimulus information as a determinant of reaction time. *J. exp. Psychol.*, **45**, 188-196.
- VINCE, MARGARET A. (1948). The intermittency of control movements and the psychological refractory period. *Brit. J. Psychol.*, **38**, 149-157.
- VINCE, MARGARET A. (1950). Some exceptions to the psychological refractory period in unskilled manual responses. *Medical Research Council Applied Psychology Unit Report No. 124/50*.
- WELFORD, A. T. (1952). The "Psychological refractory period" and the timing of high-speed performance—a review and a theory. *Brit. J. Psychol.*, **43**, 2-19.

Proceedings of the Experimental Psychology Group, 1955

29th–30th March, 1955. Extended Meeting at Oxford. *1st Session*: "New evidence on the 'Psychological Refractory Period'," by R. Davis (by invitation). *2nd Session*: "Some anomalies of perception," by M. D. Vernon. "Some differences in learning between young and adult birds in three species of finch," by M. A. Vince. "Recent observations on the mechanism of cutaneous sensibility," by A. G. M. Weddell (by invitation). *3rd Session*: "Influence of parental discipline on patterns of mental development," by N. Kent (by invitation). "The utilization of information in learning and perception," by C. E. M. Hansel (by invitation). "On figural after-effects," by F. H. George (by invitation). *4th Session*: "Some points of contact between information theory and behaviour theory," by D. E. Berlyne. "'Neutral' stimuli as operant reinforcers," by H. Hurwitz (by invitation).

12th–13th July, 1955. Extended Meeting at Cambridge, including a joint meeting on the first and second sessions with the Association for the Study of Animal Behaviour. *1st Session*: "The effects of insemination on the behaviour of *Drosophila subobscura* and some effects of inbreeding on male behaviour," by J. Maynard Smith (by invitation). "A statistical criterion for abnormal behaviour in *Anabas testudineus*," by J. B. S. Haldane and Helen Spurway (by invitation). "Displacement activities—a re-examination of the problem," by N. Tinbergen (by invitation). *2nd Session*: "'Displacement' behaviour and 'psychosomatic' disorder," by S. A. Barnett (by invitation). "The value of the concept of displacement activity in psychopathology," by D. Russell Davis. "The varieties of human conflict and frustration and their consequences," by B. M. Foss. *3rd Session*: "Errors in the visual perception of direction following a rotary body movement," by M. F. Piercy. "The use of tape recordings to study the effects of varied clarity of group goal and group path," by B. H. Raven (by invitation). Demonstrations of Apparatus by G. C. Grindley and R. L. Gregory. *4th Session*: "The influence of recall on subsequent recall," by I. M. L. Hunter. "The perception of invariance," by J. R. Beloff (by invitation).

27th–29th September, 1955. Extended Meeting at Reading. *1st Session*: "The source of sensory feed-back in the judgment of applied finger pressure," by K. A. Provins (by invitation). *2nd Session*: "Visual 'Noise'—a theory of thresholds," by R. L. Gregory. "The measurement of rhodopsin in the human eye," by F. W. Campbell (by invitation). "Some effects of brain damage on visual discrimination," by G. Ettlinger (by invitation). *3rd Session*: Short communications by members of the Departments of Psychology and Physics, Reading—R. C. Oldfield, D. J. Bruce, D. J. Gerhard and Stella Mayne and D. H. Fender. *4th Session*: Demonstrations by members of the Department of Psychology, Reading—R. C. Oldfield and J. Honisett, D. M. Vowles, D. J. Bruce, S. Laner and P. Morris, G. H. Cheesman, H. Kirkby and M. J. Townsend, and D. J. Gerhard. *5th Session*: "The effect of subliminal stimulation upon guessing habits," by N. F. Dixon (by invitation). "Effects on behaviour from noises of high and low pitch," by D. E. Broadbent. *6th Session*: "Some functions of the *Corpora Pedunculata* in the brain of ants," by D. M. Vowles (by invitation).

3rd–4th January, 1956. 9th Annual General Meeting at University College and Birkbeck College, London. *1st Session*: "Some effects of severe hypothermia on behaviour," by R. W. Russell. *2nd Session*: "Aggressive behaviour in animals following rhinencephalic lesions," by J. W. Woods (by invitation). "The utilization of uncertain cues," by J. S. Bruner (by invitation). "Some experiments on the von Restorff effect," by R. T. Green (by invitation). *3rd Session*: "Mechanisms of attention," by A. Bursill (by invitation). "Searching for a series of decisions," by N. H. and J. F. Mackworth. *4th Session*: "The mechanical recognition of speech sounds," by D. B. Fry (by invitation). "The rôle of knowledge in distance perception," by H. E. Gruber.

Committee, 1956: President: Professor R. C. Oldfield.

Editor: Dr. D. Russell Davis.

Ordinary Members of Committee: Mr. D. E. Broadbent, Mr. R. L. Gregory, Dr. W. E. Hick (Hon. Treasurer), Mr. P. H. R. James (Hon. Secretary), Dr. J. Szafran.

The following have been elected to membership of the Group.

- 1955: Berlyne, D. E., Ph.D., Department of Psychology, University of Aberdeen, Old Aberdeen.
 Crossman, E. R. F. W., M.A., Department of Engineering Production, University of Birmingham, Edgbaston, Birmingham 15.
 Elithorn, A., M.A., M.B., B.Ch., M.R.C.P., D.P.M., Neurological Research Unit, National Hospital, Queen Square, London, W.C.1.
 Hunter, I. M. L., D.Phil., Department of Psychology, University of Edinburgh, South Bridge, Edinburgh.
- 1956: Asher, H., B.A. Department of Physiology, University of Birmingham, Edgbaston Birmingham 15.
 Brown, J., Ph.D., Department of Psychology, Birkbeck College, Malet Street, London, W.C.1.
 Campbell, F. W., M.A., M.B., Ch.B., Ph.D., D.O.M.S., Department of Physiology, Downing Street, Cambridge.
 Hurwitz, H. M. B., B.A., Department of Psychology, Birkbeck College, Malet Street, London, W.C.1.
 Kelvin, R. P., Ph.D., Department of Psychology, University College, London, Gower Street, London, W.C.1.
 Uttley, A. M., Ph.D., Radar Research Establishment, Malvern, Worcs.
 Watson, A. J., B.A., Psychological Laboratory, Downing Place, Cambridge.

Resignation.

- 1955: Chambers, E. G., M.A., Psychological Laboratory, Downing Place, Cambridge.

Visiting Foreign Members.

- 1955: Baker, C. H., Ph.D., Applied Psychology Research Unit, 15, Chaucer Road, Cambridge.
 Cronbach, Prof. L. J., Ph.D., Office of Naval Research, Keysign House, 429, Oxford Street, London, W.1.
 Gibson, Mrs. E. J., Ph.D., Institute of Experimental Psychology, 34, Banbury Road, Oxford.
 Gibson, Prof. J. J., Ph.D., Institute of Experimental Psychology, 34, Banbury Road, Oxford.
 Gruber, H. E., Ph.D., Institute of Experimental Psychology, 34, Banbury Road, Oxford.
 Leeper, Prof. R. W., Ph.D., Department of Psychology, University of Aberdeen, Old Aberdeen.
 Taylor, F. V., Ph.D., Applied Psychology Research Unit, 15, Chaucer Road, Cambridge.
 Weizkrantz, L., Ph.D., University Laboratory of Physiology, Oxford.

Apparent Fluctuations of a Sensory Threshold

BY

R. C. OLDFIELD

From the Department of Psychology, University of Reading

CORRIGENDUM

In my paper "Apparent fluctuations of a sensory threshold" (this *Journal*, 1955, 7, 101-115), it was stated (page 106) that a sinusoidally modulated sine wave is physically identical with the composition of two other unmodulated sine waves such as would produce beats. This statement is, however, incorrect in principle. (It may be noted that, *mathematically*, the departure from true sinusoidal modulation in the composition of two sine oscillations may be small. The mathematical expression incorrectly used for the modulated wave, *viz.* the first of the two expressions on page 106, differs from the required form by the occurrence of $(F + f)$ instead of F in the first term. In the conditions of the experiment f/F was approximately $1/70$. This quantitative aspect, however, conceals what is acoustically essential.)

A true sinusoidally modulated sine wave is, of course, in general to be resolved into three pure sine components of frequencies F ("carrier"), $(F + f)$, and $(F - f)$ ("sidebands"). Consequently the further statement made on page 106, that the sine modulation of a pure tone might equally be perceived as the intrusion of a *fresh tone* upon the background of an existing one, should be amended to read "... of fresh tones." The subjects' reports on this experience were not in fact specific as to the exact character of the intrusion, which, at this intensity, might in any case be expected to be complex.

I am indebted to Dr. Grey Walter for drawing my attention to the erroneous nature of the original statements.

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY

Vol. VIII

1956

Part 2

SOME FACTORS AFFECTING RESPONSE IN RECOGNITION TESTS

BY

EUNICE BELBIN and VIOLET R. CANE

From the Medical Research Council Applied Psychology Unit, Cambridge

Experiments were designed to investigate some of the factors which might affect a subject's responses in a recognition test. Two experiments are described. In the first, subjects were shown 12 similar photographs. They were then given a recognition test of 12 photographs, of which a number varying from 0 to 12 had been shown in the original display. All the photographs were sufficiently alike to render some guessing essential. It was found that subjects tended to select approximately 50 per cent. as having been seen previously and to reject 50 per cent. When a second and similar display and recognition test followed immediately, there was a tendency for subjects to make more "yes" responses, without any improvement in accuracy. The effect of giving serial presentation first was to increase the number accepted not only in that test but also in a test using simultaneous presentation which succeeded it. It was also shown that when a correct decision was possible but difficult a positively correct response (in this case recognizing something seen before) was easier than a negatively correct one (rejecting something not seen before).

In the second experiment subjects were shown a picture and were then asked to describe it from memory. Immediately afterwards half the group was given a multiple-choice recognition test while the other half was given a single-choice recognition test. It was found that accuracy and certainty of decision varied with the nature of the choice situation. The results are discussed in relation to a theory of adaptation levels.

I

INTRODUCTION

When a subject has to make a choice between two (or more) responses in a situation that provides little or no information upon which to base a decision, he does not, as is well known, choose at random. If the response is one of a sequence his choice tends to be influenced by his previous response, whether this is a genuine judgement (Fernerberger, 1920; Shafter, 1950) or itself a guess (Arons & Irwin, 1932; Goodfellow, 1938; James, 1954; Karlin, 1951; Verplanck *et al.*, 1952; Whitfield, 1950; Yacorzinski, 1941). It is also influenced by the instructions of the experimenter (Senders & Sowards, 1952), the expectation of a particular result induced by the experiment as a whole (Levin, 1943), knowledge of results of previous responses (Jarvik, 1951) and previous experience of a related kind (Whitfield, 1950). If the response is isolated or the first of a set, past experience, the form of instructions and the environment of the experiment all affect it (Goodfellow, 1938; Whitfield, 1950).

In the present paper we investigate recognition-test procedures as commonly used. These tests consist of some items which have already been displayed, together with a number of similar items (controls) which have not been seen previously. The items

are presented either consecutively in book form or simultaneously on one large sheet. Although the results of attempts to validate these tests are not encouraging, very little research has been conducted on the way in which the internal features of the tests may affect a subject's performance. Postman (1950) has shown that the ratio of the number of items which have been displayed previously to the number of controls can be an important determinant of choice behaviour. He compared the effect of presenting one, three, six or nine controls with the original stimulus and found that the number of alternatives was "... a significant determinant of recognitive performance." In the present experiments we have tried to throw some light on how far and in what way recognition responses are affected by, among other things, the number of objects between which a choice has to be made and the number of those which have to be chosen. In the first experiment the number of objects out of which a choice has to be made is fixed, but the number to be chosen is not specified. The objects are shown either serially or simultaneously. In the second experiment one object has to be recognized in either a single choice or a multiple choice situation.

II

FIRST EXPERIMENT

Each of 130 subjects (Technical College students and R.A.F. personnel) was shown 24 polyphotos—12 photographs of a man (Set A) and 12 of a woman (Set B). These polyphotos were displayed simultaneously in two sets for 30 seconds. Immediately afterwards the subject was given two recognition tests, one relating to Set A and the other to Set B. In each case he was shown 12 photographs, some of which belonged to the sets shown previously and others of which were new. The new photographs varied in number from 0 to 12 for 13 different groups of subjects. Thus, for Set A photographs the first group of 10 subjects was shown 12 of the original photographs; the second group was shown 11 of the original ones and 1 new one; the third group was shown 10 originals and 2 new ones, and so on, the thirteenth group being shown none from the original display, but 12 new photographs. The procedure was the same with Set B photographs, each subject being given the same number of original photographs in Set A as in Set B. Different combinations of photographs were used for each subject in every group, as some of them might have been more easily identifiable than others. The test procedure was thus as follows:—

Display of 12 photographs Set A and 12 photographs Set B shown simultaneously.

Recognition Test.—Five of each group of 10 subjects had Set A followed by Set B; the other five had Set B followed by Set A. Set A photographs were always presented simultaneously. Set B photographs were always presented consecutively.

Results.

The number of photographs which were selected by each half group of five subjects for each condition is shown in Table I.

The maximum number that the total of any one row could reach is 240 ($= 5 \times 4 \times 12$), so that there was a tendency to choose about half the maximum number irrespective of the number of photographs actually seen before. The total, however, does appear to increase slightly as the number seen before increases. The possibility of a steady increase was investigated by including the relevant regression in the analysis of variance which was done on the data. This analysis is presented in Table II.

The effect of test and of order has been compared with the within-group residual; there appears to be no difference between showing the photographs simultaneously and showing them consecutively. Order seemed to have some effect; this is examined in detail later, since it is influenced by the interaction. The interaction (which from the form of the experiment is necessarily influenced by the differences between subjects) and the two between-group effects must be compared with the subject effect. It will

TABLE I

THE TOTAL NUMBER OF PHOTOGRAPHS WHICH WERE SELECTED BY THE FIVE SUBJECTS FOR EACH TEST CONDITION

(w) = photographs which had not actually been seen in the display
(r) = photographs which had been seen in the display

Number of original photographs present in recognition display	CONDITION 1 (Set A first, Set B second)				CONDITION 2 (Set B first, Set A second)				Total number of photographs selected as having been previously seen
	Set A		Set B		Set A		Set B		
	(w)	(r)	(w)	(r)	(w)	(r)	(w)	(r)	
0 × 5	24	—	25	—	39	—	35	—	123
1 × 5	26	0	29	2	25	3	24	4	113
2 × 5	21	8	22	5	35	8	29	7	135
3 × 5	20	10	20	12	29	9	26	11	137
4 × 5	11	11	23	11	24	13	19	17	129
5 × 5	18	19	15	18	18	14	20	14	136
6 × 5	14	15	15	17	14	17	12	18	122
7 × 5	10	21	12	24	18	29	14	27	155
8 × 5	6	26	12	23	7	28	11	26	139
9 × 5	6	26	6	27	8	29	8	29	139
10 × 5	3	19	5	29	9	37	5	34	141
11 × 5	1	24	2	35	2	37	1	35	137
12 × 5	—	42	—	42	—	29	—	42	155
	381		431		481		468		

TABLE II

ANALYSIS OF VARIANCE OF THE NUMBER ACCEPTED*

Within groups of 10 (same number seen before)	s.s.	d.f.	Mean square	v.r.
Tests (A or B)	754	13	58	not sig.
Order (1st or 2nd)	1,511	13	116.2	2.18 5 per cent.
Residual	5,567	104	53.5	
Subjects:				
Interaction Test v. Order ..	4,440	13	341.6	2.60 1 ..
Subjects within sets of five ..	13,646	104	131.2	2.45 1 ..
Total within groups	25,918	247		
Between groups:				
Regression on number seen before	1,012	1	1,012	7.7 1 ..
Residual	1,196	11	109	not sig.
	28,126	259		

Average 48.9 equivalent to 6.8 out of 12.

Slope 0.5, i.e. on the average the number accepted is increased by 0.1 when the number of photographs that have been seen before is increased by 1.

* The proportion of the number accepted out of 12 was transformed to ϕ , where $\phi = \sin^{-1}\sqrt{P}$, in order to stabilize the variance (see Kendall, 1946, p. 207).

be seen that the total number of photographs selected does in fact tend to increase linearly with the number seen before, and that there is a significant interaction, i.e. the effect of order on the two tests is different.

The fact that the number selected increases with the number seen before suggests that there is some real recognition. We have therefore estimated separately the probability of choosing a photograph seen before and the probability of choosing a photograph not seen before (Table III). The groups were pooled into three sets so as to reduce the effect of differences between subjects; the middle set is rather small because we wished to restrict it to those groups for which approximately half the test photographs were controls.

TABLE III

THE PROBABILITY OF ACCEPTING PHOTOGRAPHS NOT ACTUALLY SEEN PREVIOUSLY
= $p(w)$ AND OF THOSE WHICH HAD BEEN SEEN PREVIOUSLY = $p(r)$

Number of photo- graphs out of 12 actually seen in original display	CONDITION 1 (A first, B second)				CONDITION 2 (B first, A second)				Totals	
	Set A		Set B		Set A		Set B			
	p(w)	p(r)	p(w)	p(r)	p(w)	p(r)	p(w)	p(r)	p(w)	p(r)
0-4 ..	0.41	0.58	0.48	0.60	0.61	0.66	0.53	0.78	0.51	0.66
5-7 ..	0.47	0.61	0.47	0.66	0.56	0.66	0.51	0.66	0.50	0.65
8-12 ..	0.32	0.55	0.50	0.62	0.52	0.64	0.50	0.66	0.46	0.62
0-12 ..	0.41	0.57	0.48	0.63	0.58	0.65	0.51	0.68	0.50	0.63

The important points to be observed from Table III are noted below.

- (1) Each $p(w)$ is less than the $p(r)$ associated with it (i.e. than the one immediately succeeding it). This implies that in every experimental condition the chance of selecting a photograph seen before is greater than that of selecting one not seen before.
- (2) Although the total numbers accepted in a test done second (431 and 481) are greater than in a test done first (381 and 468) this is not due to an improvement in discrimination. Any increase in $p(r)$ between Set A done first and Set B done second is almost always accompanied by an increase in the corresponding $p(w)$, but there is no increase in any $p(r)$ and an invariable increase in $p(w)$ between Set B done first and Set A done second.
- (3) The effect of doing test B first is to increase the total number accepted in test B, and also in test A which follows, (Table I) but again this is not due to any improvement in discrimination, since for Condition 2 as a whole $p(w) = 0.55$, $p(r) = 0.67$; while for Condition 1 $p(w) = 0.45$ and $p(r) = 0.60$.
- (4) We see from the totals that when all four tests are pooled both $p(w)$ and $p(r)$ decrease as the number of photographs seen before increases.

This last point suggests that the subjects are partly adapting their behaviour in such a way as to select a fairly constant number of the 12 photographs shown, although some real recognition is still influencing their results. Table IV helps to clarify this point. The first column shows the number of photographs that would be accepted in each of the three cases if $p(w)$ and $p(r)$ were 0.50 and 0.63 in each case, i.e. if the overall proportions of acceptance both wrongly and rightly remained constant despite changes in the proportion of original photographs present. The second column shows

the number that would be accepted if the overall probability of acceptance remained constant at 0.565 (the average of 0.50 and 0.63), i.e. if subjects chose always the same proportion of the photographs in the test. The third column shows the observed results.

TABLE IV

Number of photographs out of 12 actually seen in original display	Number that would be accepted if—		Number actually accepted
	Ratio $\frac{p(w)}{p(r)}$ Constant	$p(w) + p(r)$ Constant	
0-4	625	677	637
5-7	407	407	413
8-12	729	677	711
	1,761	1,761	1,761

It can thus be seen that the number actually accepted represents on the whole a compromise between Columns 1 and 2.

There seems to be some confirmation of such adaptation of behaviour from the following figures (Table V) which show for Test B the number of photographs selected from the first six in the series and from the last six in the series.

TABLE V

THE NUMBER OF PHOTOGRAPHS SELECTED FROM THE FIRST SIX IN THE SERIES AND FROM THE LAST SIX IN THE SERIES (SET B)

Number of control items out of 12	Proportion of first six chosen	Proportion of last six chosen
8-12	149/300 = 0.50	172/300 = 0.57
5-7	103/180 = 0.57	103/180 = 0.57
0-4	187/300 = 0.62	185/300 = 0.62

Since the photographs were shown in random order we should expect the proportion selected to be the same for the first six as for the second six in each row of the table if $p(w)$ and $p(r)$ remained the same throughout the series, and this in fact is found in the last two rows of the table. When the number of control items is high, however, the subject behaves as though he had not chosen enough at the beginning and had to compensate by choosing more towards the end.

Seventy-one subjects reported that they found the simultaneous presentation of Set A easier than the consecutive presentation of Set B; 38 preferred B and 21 had no preference.

Discussion

The results of this experiment suggest that the number chosen in a recognition test may be determined by the total number of items presented. The present experiment was designed with an arbitrary total of 12 items. Similar trends were found in a pilot experiment with a total of 10 items. It is of importance to note, however, that the present experiment used the same total number of items in the

original display of photographs as were used in the recognition test. The extent to which the relationship between number originally presented and number presented in the test affects the results is a matter for further experimentation.

Whitfield (1950) has suggested from his experiment with an imaginary questionnaire that where doubt exists in a test the number of "yes" responses is greater than 50 per cent. of the total. In the present experiment, however, it was shown that the number of "yes" responses in excess of 50 per cent. of the total was proportionate to the actual number of items that had been seen before. A further comparison with his results is of interest. It will be seen from Table V that the percentage of initial "yes" responses in his experiment well exceeded the percentage in the present experiment.

TABLE V
PERCENTAGE OF INITIAL "YES" RESPONSES

	<i>Present experiment, 12 questions</i>	<i>Imaginary questionnaire</i>	
		<i>10 questions</i>	<i>14 questions</i>
Percentage of initial "yes" responses to the total initial responses . .	63	87.5	82.5

Whitfield in his experiment used test-sophisticated subjects. The subjects in the present experiment were unlikely to have been familiar with the procedure of recognition tests. The effect of previous experience on test results is suggested by the significant difference which was found in the present experiment between tests done first and tests done second. Thus in any further experiments which are conducted to investigate the nature of responses in a choice situation it seems essential that subject differences should be taken into account. This need is further emphasized in an investigation by R. M. Belbin (unpublished paper), who used a recognition test as a criterion of amount learned by older and younger people. He noted that there was a significant difference in the "yes" responses in recognition tests between people over 50 and those under 50, those over 50 tending to accept doubtful items much more often than the younger subjects.

III

SECOND EXPERIMENT

The following experiment was designed to assess to what degree and how a multiple-choice test as compared with a single-choice test may affect a subject's final decision in identification.

Experimental Design

A picture was shown to each subject for 5 seconds. He was then asked to describe as much of it as he could from memory and was immediately afterwards given a recognition test. The 60 subjects (research students and naval ratings) were divided into two groups as follows:—

Group A was given a multiple-choice recognition test. They were asked to say which of six pictures they had seen previously. The pictures differed from each other in one detail only.

Group B subjects were given a single-choice recognition test. They were shown the picture which they had originally seen and were asked to say whether or not it was the original.

Each subject was asked to give his response on a three-point scale of certainty.

RESULTS

Table VI shows the number and degree of certainty of responses of the two groups.

TABLE VI

	Number of subjects giving correct responses			Number of subjects giving incorrect responses		
	<i>Very Sure</i>	<i>Sure</i>	<i>Doubtful</i>	<i>Doubtful</i>	<i>Sure</i>	<i>Very Sure</i>
Group A	1	3	3	7	8	8
Group B	3	4	8	9	6	0

Incorrect responses were scored as follows:—

For Group A—when a subject chose a picture other than the original.

For Group B—when a subject said "No, that is not the original."

There appears to be a difference in quality of response between Groups A and B in that there were nine responses of which the subject was very sure in Group A and only three in Group B; this difference is not, however, significant. The number of correct responses appears very different in the two cases, but it is not really possible to compare them since it is not clear how far random guessing is involved. For Group A we can obtain a corrected score as Postman does, but the method cannot be applied to Group B. The reason for this is perhaps worth considering since the correction has been applied in the single choice case by, for instance, Votaw (1936).

Let P be the probability of truly recognizing the picture and consequently of accepting it; then $1 - P$ is the probability of not recognizing it. Let $p(w)$ be the probability of choosing a picture that is not recognized. Then the total probability of choosing the correct picture is $p(r)$, say, where

$$p(r) = P + (1 - P)p(w)$$

Write E for the average over the population of subjects. Then

$$E(p(r)) = E(P) + E(p(w)) - E(Pp(w))$$

In the multiple choice experiment it is reasonable to assume that $p(w) = 1/6$, since each alternative picture differs in only one detail from the true one and so one may suppose that each of the six is equally likely to be selected when there is no true recognition. In this case P and $p(w)$ will be independent and the equation above becomes

$$E(p(r)) = E(P) + 1/6 - 1/6E(P)$$

$$\therefore E(P) = \frac{E(p(r)) - 1/6}{1 - 1/6}$$

$E(p(r))$ may be estimated by the average number right per subject, in this case $7/30$,

so that $E(P) = \frac{7/30 - 1/6}{1 - 1/6} = 2/25 = 0.08$ and this is the average corrected score.

Votaw uses a formula similar to the last in the single choice case, taking $p(w) = \frac{1}{2}$; this, however, assumes not only that on the average a subject is equally likely to say yes or no when he does not know the answer, but also that his chance of saying yes or no is independent of his power of giving the true answer. The results of Whitfield *et al.* contradict the first assumption, and there appears to be no experimental evidence the one way or the other with regard to the second. Thus the apparent simplicity of the

single choice method is misleading. It is possible, of course, to estimate $p(w)$ for each subject by giving him single choice tests in which he is shown one of the alternative items, not the one previously seen; but such testings would have to be separated in time since the data of the previous experiment suggest that in a serial presentation successive responses are not independent.

IV

DISCUSSION

In tests of recognition subjects are asked to discriminate between items which have and have not been exposed previously. Zangwill (1939) has shown how outstanding detail provides the principal clue in identifying material under experimental conditions. In these experiments, however, as in most recognition tests used in the laboratory, no one item had any detail which made it outstandingly different from any other item. Thus, when attempting to match the items with the image of the original, identification was likely to rest only on a feeling of familiarity—or more probably on a feeling of unfamiliarity. Since only barely discriminable differences separated the items none of them were likely to be totally unfamiliar. The subject must gradually have begun to realize that identification was determined by the items being more or less familiar. He had therefore to provide himself with a scale or level of relative familiarity. Since his image of the original will have been affected by the interpolated activity of looking at other similar items, this scale or level will have changed schematically as he progressed through the test.

It would appear then that the process of recognition in this type of situation is a matter of making a judgement very much akin to the type of judgement required in psychophysical experiments on thresholds and a phenomenon similar to the "central tendency" or "level of adaptation" may well determine identification.

Michels and Helson (1949) define the adaptation level, A.L., as that level which evokes a neutral response. (Thus, in judging weights, a stimulus above the A.L. is judged as heavy, one below the A.L. as light.) They suggest that this A.L. is determined not only by the stimuli judged, but by all other stimuli present as a background and residual stimuli from past experience. They appear to consider that the stimuli can be correctly judged, that is, correctly ranged on some scale, but that the response given varies because the point of reference, the A.L., itself continually varies as the experiment proceeds. They imply that the A.L. is determined by the subject's need to make his scale of judgements cover the range of stimuli considered in some appropriate way. Thus the A.L. may be regarded as determined by the range of the stimuli so far seen and the subject's idea of the frequency with which his responses should come into the various judgement categories.

In recognition tests of the type which we have considered there is, of course, no absolute measure that can be applied to the stimuli. We can, however, in the simple case of our second experiment think of the stimuli as being ordered on a scale of which the left-hand end represents "same" and points to the right of this represent various degrees of difference. In the more complicated situation of the first experiment we have to think of twelve parallel scales all considered at once, one for each of the original photographs. The position of the A.L. is represented by the number of objects accepted.

This interpretation of the problem as one of scaling the items in the tests seems to be borne out by the subjects' own comments. Unless a subject recognized the correct item immediately it seemed that he collected all possible similarities between his recall of the original and the other items and proceeded by a method of rejecting the

unfamiliar. Practically every subject in the second experiment made his final selection in the test by eliminating what he thought were unlikely items and by accepting implicitly the fact that it must be one of the items. Only three of the subjects showed any doubt at all that the original was actually present.

It seems then that experiments of this type depend, apart from the material used, on the range of stimuli used in the experiment and the fineness of discrimination required (which may be determined by instructions or simply by the subject's own ideas of what is fitting). Similar variables must also influence other choice situations where an element of doubt is present, such as, for example, responses in intelligence tests, in parapsychological experiments or in time-study ratings. There exists as yet no theory by which the effects of these variables can be predicted. Overall formulae used in recognition tests to compensate for the "error of falsification" (i.e. the assessment of the number of items which are likely to be accepted erroneously) are, in our present state of knowledge, quite meaningless.

BIBLIOGRAPHY

- ARONS, L. and IRWIN, F. W. (1932). Equal weights and psychological judgements. *J. exp. Psychol.*, **15**, 733-751.
- FERNBERGER, S. W. (1920). Interdependence of judgements within the series for the method of constant stimuli. *J. exp. Psychol.*, **3**, 126-150.
- GOODFELLOW, L. D. (1938). Zenith radio experiments in telepathy. *J. exp. Psychol.*, **23**, 601-632.
- JARVIK, M. E. (1951). Probability learning and a negative recency effect in the serial anticipation of alternative symbols. *J. exp. Psychol.*, **41**, 291-297.
- KARLIN, J. (1951). Influence of equality judgements on the constant error. *J. exp. Psychol.*, **42**, 300-303.
- KENDALL, M. G. (1946). *Advanced Theory of Statistics*, Vol. II. London.
- LEWIN, M. M. (1943). Weight illusions induced by cues of low validity. *Psychol. Bull.*, **40**, 582.
- MICHEL, W. C. and HELSON, H. (1949). A reformulation of the Fechner Law in terms of adaptation-level applied to rating scale data. *Amer. J. Psychol.*, **62**, 355-368.
- POSTMAN, L. (1950). Choice behaviour and the process of recognition. *Amer. J. Psychol.*, **63**, 576-583.
- SENDERS, V. L. and SOWARDS, A. S. (1952). Analysis of response sequences in the setting of a psycho-physical experiment. *Amer. J. Psychol.*, **65**, 358-374.
- SCHAFER, T. H. (1950). Influence of preceding item on measurement of noise. *J. exp. Psychol.*, **40**, 365-371.
- VERPLANCK, W. S., COLLIER, G. H. and COLTON, J. W. (1952). Non-independence of successive responses in measurement of the visual threshold. *J. exp. Psychol.*, **44**, 273-282.
- VOTAW, D. F. (1936). The effect of do-not-guess directions on true-false or multiple choice tests. *J. educ. Psychol.*, **27**, 698-703.
- WHITFIELD, J. W. (1950). The imaginary questionnaire. *Quart. J. exp. Psychol.*, **2**, 76-87.
- YACORZINSKI, G. K. (1941). Perceptual principles involved in the discrimination of a configuration formed in predicting the occurrence of patterns selected by chance. *J. exp. Psychol.*, **29**, 401-406.
- ZANGWILL, O. L. (1939). Some relations between reproducing and recognizing prose material. *Brit. J. Psychol.*, **29**, 370-382.

SOME EFFECTS OF THIAMINE DEFICIENCY AND REDUCED CALORIC INTAKE ON "BEHAVIOUR UNDER STRESS" AND ON LEARNING

BY

F. KNOPFELMACHER, MELEK KHAIRY, R. W. RUSSELL AND JOHN YUDKIN

From the Department of Psychology, University College London, and the Department of Nutrition, Queen Elizabeth College, University of London

The main purpose of the investigation was to determine the effects of thiamine deficiency in rats on behaviour under stress and on learning. Since thiamine deficiency reduces food intake, it was necessary to have two control groups. The first received adequate thiamine, but a food intake reduced to that of the vitamin-deficient animals; the second received adequate thiamine and an unrestricted food intake. It was thus possible to study the effects both of thiamine deficiency and of reduced caloric intake on the behaviour variables measured.

The experimental group was maintained on a thiamine-deficient diet throughout the entire experimental period. Because of the effects of thiamine deficiency on caloric intake, food-hunger was never used as a form of motivation. Behaviour in four different situations was studied: in Hall's open-field test, in two discrimination situations both involving exposure to insoluble and soluble problems, and in a water maze.

Contrary to the findings of several previous studies, there were no significant effects of thiamine deficiency on behaviour prior to the onset of polyneuritis with its debilitating effects on motor co-ordination. There was also no evidence of impairment in any of the behaviour studied which could be attributed to restrictions in caloric intake. This was the case even though the restrictions continued for a period of somewhat more than twelve months. Although these restrictions did not lead to impairment they were associated with certain changes in performance. Animals whose feeding was restricted were more active and, during the soluble phase of one discrimination situation, showed more vicarious trial and error and learned more readily than animals fed *ad libitum* on the same diet. It is suggested that these differences may be interpreted in terms of the effects of what previous investigators have referred to as "irrelevant drives."

I

INTRODUCTION

Systematic investigations of the behavioural effects of nutritional deficiencies are important for at least three reasons. First, there is the purely practical necessity of obtaining information on possible ways in which various types of nutritional deficiency affect the behaviour of an organism. Second, the induction of nutritional deficiencies may be employed to obtain additional information on the relations between neurophysiological processes and behaviour, since changes in central nervous system activity may in some cases be inferred from the resultant behaviour patterns, and, conversely, if the biochemical concomitants of the deficiency are known, they may be used as agents for producing "biochemical lesions" in the nervous system. Thirdly, in studies of the learning process food deprivation is widely used to motivate experimental subjects. It is well known that the generality of laws postulated on the basis of such studies is often restricted to one type of learning situation and the results are frequently not reproducible. It may be that the reason for this lies partly in the very general nature of the commonly used "drive" concept. Deprivation could be conceived to precipitate behaviour patterns which may not only vary with duration, onset, and the time-intervals between it and behavioural tests, but also with a great variety of

biochemical processes affected by specific deficiencies. It is possible that "specific hungers" may result in specific behaviour patterns which would depend on the complex inter-relations between the physiological processes involved and the types of behaviour investigated.

Studies of behavioural effects of thiamine deficiency are particularly interesting in view of the effect thiamine reduction has on the carbohydrate metabolism of the brain. Thiamine deficiency affects the normal oxidative mechanism of the brain and it has been claimed that such effects impair the learning ability of experimental animals. Some results have been obtained which purport to confirm this hypothesis (e.g. Maurer, 1935; Stevens, 1937; Poe, Poe & Muenzinger, 1937; Biel & Wickens, 1941). It appears from these reports that at least three variables are important with regard to the effects of thiamine deficiency on behaviour: the age at which the deficiency begins, the duration of the deficiency, and the length of the time interval, if any, between supplementation of previously deprived animals and the beginning of experimental observations of behaviour. These three variables have not been controlled systematically in previous behaviour studies. Frequently, food deprivation has been used to motivate thiamine deficient animals in experimental situations; the known effects of thiamine deficiency on caloric intake, with anorexia as one of the symptoms, makes such a procedure dubious. Other controls and checks, e.g. on growth rates and food intake, have also been neglected.

The present study was designed to determine whether or not impairment in behaviour would occur when these various matters were taken into account. Behaviour was observed in four different experimental situations. The three used techniques for exposing the subjects, male hooded rats, to "stress" (Russell, 1953), techniques which have been used extensively for research in experimental psychopathology. Two of these provided measures of behaviour during and after exposure to stress. The fourth situation yielded conventional maze-learning scores.

II

METHOD

The experimental design is shown in Table 1. It required three groups of animals: a general control group, a thiamine-deficient group and a caloric control group. The latter served as a control for the effects of reduced caloric intake in the thiamine-deficient group. Only the general control group received an *ad libitum* supply of all the necessary dietary components.

TABLE 1
EXPERIMENTAL DESIGN

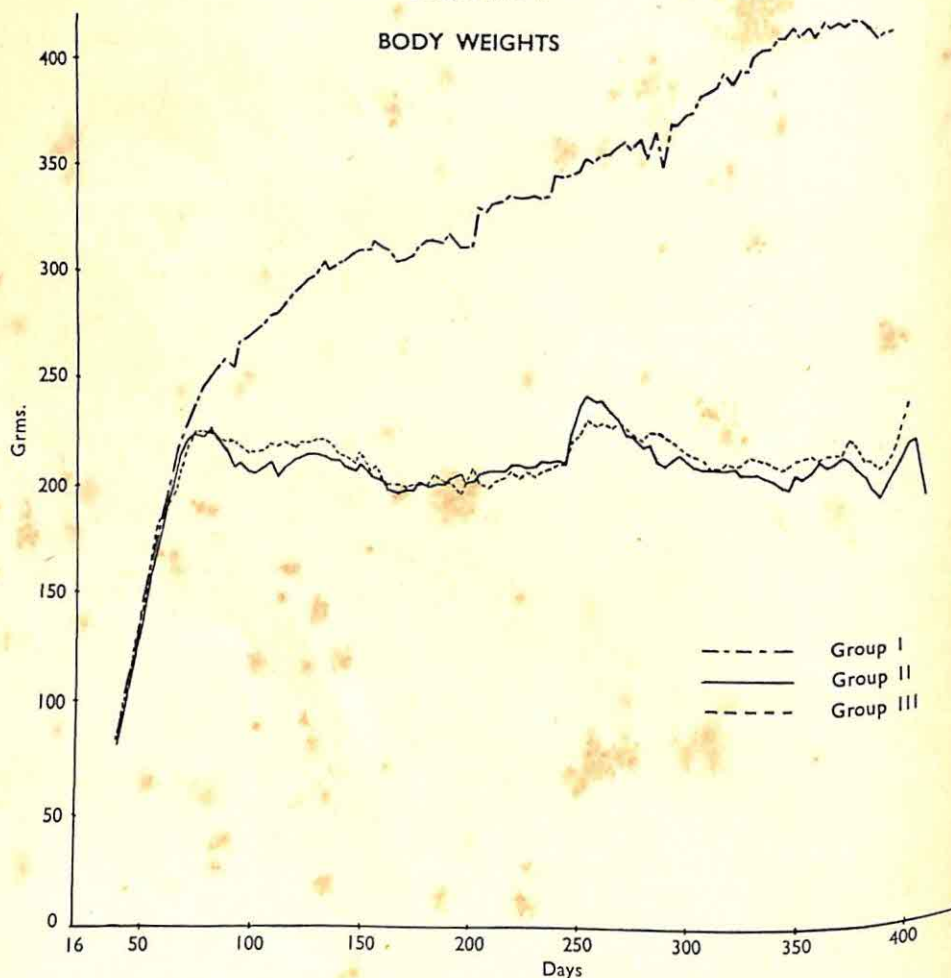
Group	Diets	Phases			
		1	2	3	4
I	Thiamine-supplemented <i>ad libitum</i>	Open-field test	Insoluble problem 1	Insoluble problem 2	Water-maze learning
II	Thiamine-deficient <i>ad libitum</i>	Open-field test	Insoluble problem 1	Insoluble problem 2	Water-maze learning
III	Thiamine-supplemented caloric control	Open-field test	Insoluble problem 1	Insoluble problem 2	Water-maze learning

Subjects.

Twenty-one male hooded rats of a highly inbred Lister stock from the animal colony of the Department of Nutrition, Queen Elizabeth College, University of London, served as subjects. Males from five litters weaned at 23 days of age were selected, and were divided into seven litter-mate trios by weight. Between 49 and 51 days of age the animals were assigned to three groups in accordance with customary litter-mate-control techniques. At this time they were started on the feeding schedules and diets appropriate for the groups to which they belonged and they continued on these diets until the investigation was completed.

A general routine for handling the animals was followed daily throughout the experiment. In the morning they were weighed and fed thiamine supplements where required, and the food left in their feeding jars was also weighed. In the afternoon the animals were taken individually from their cages, given the appropriate trials for the day, and returned to their cages.

FIGURE 1
BODY WEIGHTS



The curves represent the median in grams of the body weights recorded per day for animals in each group.

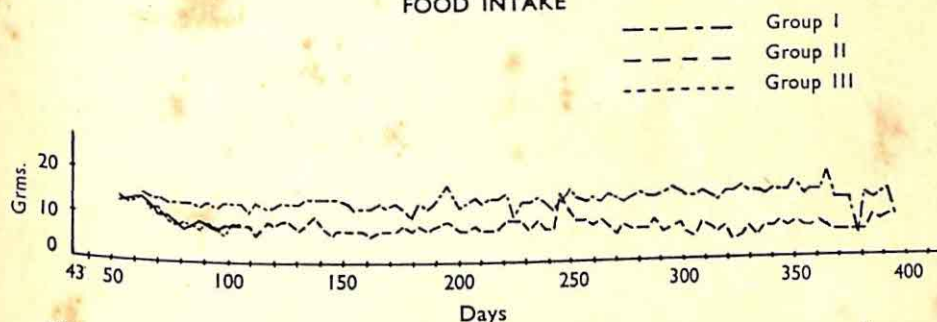
Diet and Growth.

Table II shows the composition of the diet and of the supplements. Group I, the general control group, was allowed an *ad libitum* food intake, supplemented with adequate doses of thiamine. Group II was maintained on a thiamine-deficient diet. For these animals the daily intake of thiamine was 5 μ g., an amount which previous experience

suggested would give suboptimal growth, but would not lead to polyneuritis or death for at least several months (Yudkin, 1949). Animals in Group III received daily the amount of food consumed on the preceding day by their respective litter-mates in Group II and their diet was supplemented with adequate quantities of thiamine.

The animals were weighed five times a week and their food intake was checked daily to the nearest gramme. After 19 days on the thiamine deficient diet, growth rates of animals in Groups II and III became significantly less ($p = < 0.05$) than those of Group I. This differentiation of the groups is shown in Figure 1, which presents the median-weight curves for the entire period of the experiment. The food intake of Group II and, by experimental artifact, also of Group III dropped considerably below the corresponding intake of Group I when the growth rates became differentiated. Curves for median food intake are plotted in Figure 2. Observations of behaviour in the first experimental situation began after growth rates had continued to be significantly different for three consecutive days

FIGURE 2
FOOD INTAKE



The upper curve represents the median in grms. of the food consumed per day by Group I. The lower curve represents the median in grms. of the food consumed per day by Group II. The dark curve represents the median in grms. of the food given to Group III. This curve was discontinued after 105 days since Group III consumed the whole of the food given to them (identical to that consumed by Group II).

It is important to emphasize that food-hunger was never used as a form of motivation in the present research. The reasons for this are apparent and arise primarily from, first, the rigorous dietary controls required by the experimental design and, second, the effects of thiamine deficiency, e.g. anorexia, on caloric intake.

TABLE II

	Diet	
Sucrose	600	gm.
Low vitamin casein	200	gm.
Arachis oil	150	gm.
Salt mixture	50	gm.
Dry vitamins (DV ₂)	0.5392	gm. (see DV ₂ below)
Choline chloride	1.0	gm.
L-Cystine	2.0	gm.
Vitamin B ₁₂	100	µg.

Dry Vitamins (DV₂) per kg. diet:

Inositol	200	mgm.
Nicotinic acid	100	mgm.
Ca-D-pantothenate	100	mgm.
Para-aminobenzoic acid	75	mgm.
Riboflavin	30	mgm.
Pyridoxin	8	mgm.
Biotin	0.2	mgm.
Folic acid	1.0	mgm.

SUPPLEMENTS

1. Thiamine (Vitamin B₁). Each animal in groups I and III received daily 20 µg. of thiamine diluted in 0.5 ml. N/500 HCl containing 25 per cent. alcohol. Animals in group II received daily 5 µg. of thiamine diluted in 0.5 ml. alcoholic N/500 HCl.
2. 100 mg. C.L.O. containing 400 µg. of Vitamin K were given twice weekly to all animals.
3. 2 mg. α-tocopheryl acetate were given once a week to all animals.

Open-field Test.

This experimental situation has been frequently employed (Hall & Ballachby, 1932; Hall, 1936; Billingslea, 1942) to assess differences in what Hall believes to be the emotionality syndrome of the rat. A circular field, 6 feet in diameter, was zoned off by a fence 18 inches high and illuminated by a 200-watt electric bulb suspended over the centre. The floor of the field was divided into segments by drawing concentric circles and lines, thus enabling an observer to record a subject's movements during each trial. The animals, now 70-72 days old, were placed singly in the field and left there for 2 minutes. This procedure was repeated on twelve consecutive days and the following observations were recorded: amount of locomotion and frequency of urination, defecation, rearing and preening. Hall and others (e.g. Billingslea, 1942) have claimed that high rates of defecation and urination together with inhibited locomotion are correlated and that they form a reponse syndrome, tentatively labelled "emotionality," which is genetically determined. Rearing and preening have been referred to by other investigators as "displacement activities." Basically this apparatus provides a mildly stressful situation in which avoidance reactions are produced by flooding the field with the bright light and actual avoidance prevented by the restraining walls of the field. In Hall's procedure the animals are usually kept on a hunger schedule and food placed in the centre of the field under the lamp. For reasons already given this aspect of the procedure had to be omitted in the present study.

Insoluble Problem 1.

Ten days after completion of the open-field test observations were begun in the first insoluble problem situation. The animals were then 91-93 days old. The procedure has been used extensively (Maier, 1949) in studies of "abnormal fixation" and other forms of aberrant behaviour. Essentially it consisted of two phases. During the first, observations were made of the animals' reactions when forced to respond to an insoluble discrimination problem and, during the second, the after-effects of this treatment were studied in terms of the animals' behaviour in a soluble discrimination situation. The apparatus, originally designed by Lawrence (1949), was in all essentials analogous to the familiar Lashley jumping stand. Since it was impracticable to use food as an incentive, the animals were forced to respond on each trial by being gently pushed from the starting box. On being pushed they had to jump to one of two landing platforms, either of which could be locked in position and therefore be "safe." If an animal jumped to the unlocked platform, it was punished by falling into a net slung below. After preliminary training each animal was presented with an insoluble problem by varying the side of the safe platform in a random order. The nature of the problem was such that the animals were punished on approximately 50 per cent. of their trials. Ten trials a day were given for ten days.

After these 100 trials the problem was made soluble by associating the safe platform and the corresponding side of the supporting frame with black cards, thus differentiating it from the white of the unlocked platform. The animals were given ten trials a day for 30 days, at which time none of them had shown any sign of learning and this phase of the experiment was discontinued.

Insoluble Problem 2.

Sixteen days after completion of the previous observations trials began in the third experimental situation, the animals being 152-154 days old. This situation consisted of an insoluble discrimination problem similar in general characteristics to the one just described, but differing in mode of response and in means of inducing a response. The apparatus and technique have been described in detail elsewhere (Russell & Pretty, 1951; Knopfmacher, 1953). The animals had to swim the length of a straightway in order to reach two doorways through which they could escape from the water. In each doorway a door was hinged in such a fashion that it could be locked or unlocked by the experimenter.

An insoluble problem could be presented by locking and unlocking the doors in a random sequence. A response to the unlocked door was reinforced by immediate escape from the water, while a response to the locked door was punished by delaying escape until the animal swam round a partition between the doorways and pushed through the unlocked door.

After preliminary training, the animals were given 100 trials on the insoluble problem at the rate of five successive trials per day. The problem was then altered so as to be soluble by always placing a black panel above the unlocked door and a white panel above the locked door. The position of the unlocked door was varied from one doorway to the other in a random sequence. The animals were given five trials each day and the learning criterion was reached when an animal swam to the unlocked door during all trials on two successive days. Some animals with strong position stereotypes did not reach the learning criterion within 400 trials and, thus, continued to be punished on approximately 50 per cent. of the trials. For them the procedure was changed so that the black panel and the unlocked door were always on the non-stereotyped side, thus associating the stereotyped response with 100 per cent. punishment. The stereotype was considered to be broken when an animal made ten correct responses and the random sequence for locking and unlocking the doors was then resumed. If new position stereotypes were set up, this procedure for breaking them was re-introduced. Even under this procedure a few animals failed to reach the learning criterion within 300 trials and were then given "guidance" using the technique suggested by Maier (1949). This consisted of guiding them to the unlocked side whenever they started to make an incorrect response. The slowest learner reached the learning criterion within 140 trials after this technique had been introduced.

Lashley Number III Maze.

Certain results obtained in the water discrimination situation suggested the desirability of further tests for variability of learning among the three groups. Lashley's Number III Maze adapted for use with escape from water as the form of motivation (Braun, Russell & Patton, 1949) was employed. The animals were now 391-393 days old. After preliminary training in a straightway, they were required to swim through the maze, one trial being given per day. The learning criterion was reached when an animal swam through the maze without error on two consecutive days.

III

RESULTS

In analysing the results for significances of differences among the various groups Mood's (1950) distribution-free median test was used. "Significance" is defined in terms of the 5 per cent. level of confidence.

Open-field Test.

The measures of performance recorded were amount of locomotion and frequency of urination, defecation, rearing and preening. None of these measures showed differences which could be attributed to the effects of the thiamine deficiency. However, animals on limited caloric intake did show significantly greater ($p < 0.01$) amounts of locomotion than those fed *ad libitum*.

Insoluble Problem 1.

Exposure to the stress of insoluble problem situations of the kind employed in the present research customarily leads to the establishment of stereotyped behaviour patterns (e.g. Maier, 1949; Russell & Pretty, 1951). Observation of our animals showed very clearly that the stereotyped behaviour took the form of a position response, i.e. the animals tended to respond during each trial by jumping to the same platform regardless of their experiences on previous trials. The following technique was used in order to arrive at an estimate of the consistency of these stereotyped behaviour patterns. The 100 responses in the insoluble problem situation were divided into sets of ten and for each set the number of the preponderating response

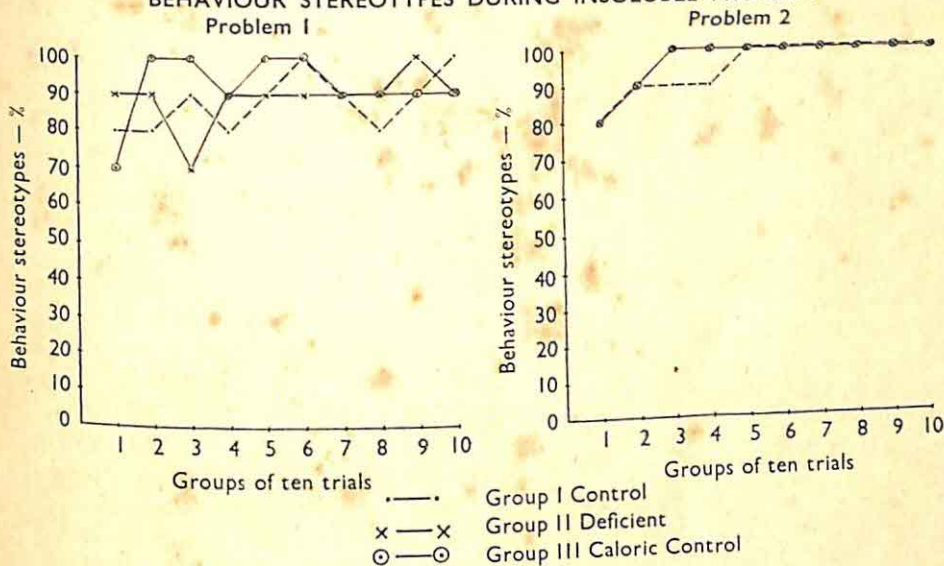
was noted. For instance, jumping three times to the left and seven times to the right was recorded as "7." The reverse, three jumps to the right and seven to the left, was also recorded as "7." When the number of left and right responses was equal, "5" was assigned. For each animal and for each set of ten trials a score was obtained, representing the mean consistency with which an animal responded. The graph on the left in Figure 3 summarizes the results of this analysis. It will be seen that highly stereotyped behaviour patterns developed very early during exposure to insoluble problem 1, but they never reached a continued level of 100 per cent. consistency. Analysis of the median scores for the total 100 trials showed no significant difference among the three groups of subjects.

Abortive jumping, i.e. jumping in some direction other than to the landing platforms, occurred occasionally. Analysis of this form of behaviour also revealed no significant differences among the groups.

It is customary in using the present technique to study the after-effects of exposure to the insoluble problem in terms of measures of performance in a new but similar soluble problem situation. In the present instance this situation consisted of a problem the solution of which depended upon learning to discriminate between white and black cards. None of the animals learned this discrimination in 300 trials, at which point practice was discontinued. It was possible that, despite this failure to reach any reasonable criterion of learning during the 300 trials, the animals might have shown certain signs of learning on the basis of which the groups could be distinguished. One of these signs in discrimination learning of the kind involved in the present research is well known as "vicarious trial and error," or "VTE" (Muenzinger, 1938). As Munn (1950) has pointed out, "It is quite apparent . . . that until such behaviour appears there is little apparent progress toward solution of discrimination problems." VTE is of particular importance in studies of the present kind, since it is one of the first signs that behaviour stereotypes are being broken, a necessary antecedent to the learning of the soluble problem. Analysis of the frequencies with which VTE occurred during the soluble phase of this first discrimination situation revealed no significant differences among the three groups of animals being tested. There were also no significant differences among the groups in regard to their median percentages of correct responses. The performances of these animals contrasted very strikingly with normal control animals, run in previous experiments on the same apparatus, but with food-hunger motivation and no exposure to the insoluble problem, who reached the criterion of ten successive runs without error in an average of 56.30 trials and with a range of 20 to 100 trials (James, 1954).

Insoluble Problem 2.

The second insoluble problem situation was similar to the first in providing a situation for the establishment of position stereotypes, but differed in that escape from water was used to motivate responses. The graph on the right in Figure 3 summarizes the results of analysing for consistencies of behaviour applying the same technique as used in describing the corresponding data from insoluble problem 1. It will be seen again that highly stereotyped behaviour patterns developed very early. Although the normal control group took somewhat longer to reach 100 per cent. consistency than the other two groups, it is interesting to note that, once the 100 per cent. level was reached, there were no deviations from it in any of the groups. This differed from the corresponding performance in the first insoluble problem situation where variations from this level occurred frequently. Analysis of the median scores for the total 100 trials again showed no significant differences among the three groups of subjects.

FIGURE 3
BEHAVIOUR STEREOTYPES DURING INSOLUBLE PROBLEM

On the other hand, differences among the groups were clearly apparent when behaviour during the soluble phase was analysed. Table III summarizes the performances of the three groups, first in terms of the numbers of animals above and below the median scores for all animals on measures of VTE and, second, the numbers of animals learning under 50 per cent. punishment, and under 100 per cent. punishment and guidance. There were significant differences among the groups on these two measures, but the differences were not between the thiamine-deprived and the two control groups. It was the caloric control group which showed signs of breaking its stereotyped behaviour patterns and of learning the soluble problem more rapidly than the other groups. These were the animals who were on a restricted caloric intake while the other animals were feeding *ad libitum* on their particular diets. All the caloric control animals were above the general median for the three groups in amount of VTE and were significantly different from the normal control animals ($p < 0.01$). This difference in VTE was reflected, as would be expected, in differences ($p < 0.05$) between the caloric control and the other two groups in measures of learning. None of the latter learned the soluble discrimination problem until the level of punishment associated with their stereotyped responses was increased from 50 to 100 per cent. and some had to be given "guidance" before the criterion of learning was reached.

TABLE III
INSOLUBLE PROBLEM 2: SOLUBLE PHASE

		Group		
Behaviour Measure		I	II	III
VTE ¹	Above median	1	3	7
	Below median	6	4	0
	Learning under 50 per cent. punishment	0	0	4
Learning ²	Learning under 100 per cent. punishment and guidance	7	7	3

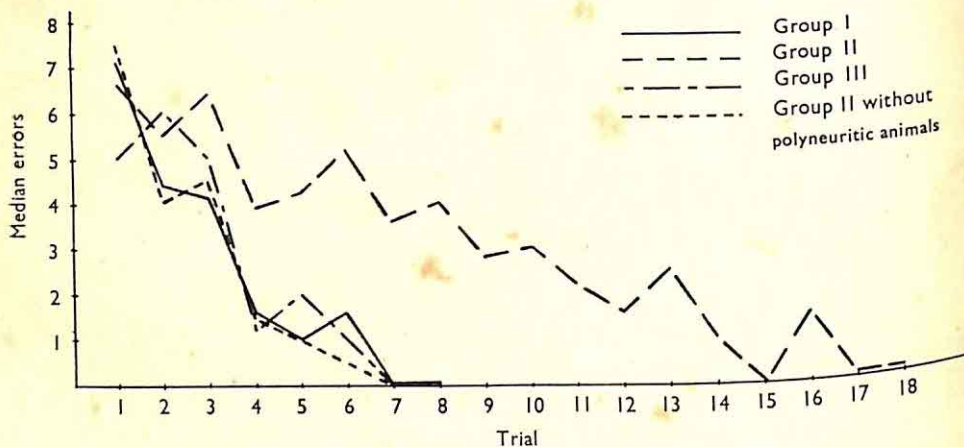
¹ The differences between Groups I and III are significant ($p < 0.01$).

² Differences between Groups I and III, and between Groups II and III are significant ($p < 0.05$).

Maze Learning.

The number of trials to reach the learning criterion and the total number of errors made during these trials served as the measures of performance in the water maze. Figure 4 presents learning curves for the three groups. It is clear from these curves that the thiamine-deficient animals (Group II) as a group, were behaving quite differently from animals in the two control groups. During the learning all but two of the animals in Group II began to show the signs of muscle rigidity and motor inco-ordination associated with polyneuritis. One of the affected animals died and the locomotion of the others was seriously impeded. The degree of the impediment varied from day to day and was associated with an increased scatter of error scores.

FIGURE 4
LASHLEY III MAZE. ERRORS DURING LEARNING



As shown in Figure 4, when the polyneuritic animals are excluded from Group II, differences between it and the other groups disappear. This observation is supported by statistical tests for the significances of the differences.

IV

DISCUSSION

Three main points of interest arise from these results.

Effects of Thiamine Deficiency.

There is no evidence from our results that thiamine deficiency had any significant effects on the types of behaviour studied, i.e. behaviour under stress and learning performance, prior to the onset of polyneuritis with its debilitating effects on muscle tonus and motor co-ordination. The discrepancy between these results and results obtained by previous investigators may arise from several differences in the research techniques employed.

First, in some studies (Maurer, 1935; Stevens, 1937) involving maze learning, food was used as an incentive and adequate allowance was not made for the effects of thiamine deficiency on motivation. Differences in learning scores which appeared to favour control groups may have been an artifact produced by anorexia among the thiamine deficient animals.

Second, certain studies (Poe, Poe & Muenzinger, 1932; Biel & Wickens, 1941), although using some form of motivation and reinforcement other than food-hunger, included no separate controls to check specifically on possible effects of general caloric

deficiency. Information about such effects would seem to be important for the interpretation of any observed differences in behaviour between dietary deficient and normal control animals.

Third, the majority of investigators kept their animals on a thiamine-deficient diet by way of depleting the lactating mother and supplementing the animal's diet before experimental observations began. It is possible that some of the results obtained using this technique may have been due to effects produced in early infancy and associated indirectly with learning ability. These effects may have been due either to thiamine or to general caloric deficiency.

Fourth, with one exception (Stevens, 1937), investigators appear not to have controlled adequately the degree of thiamine deprivation. Therefore, it was difficult to distinguish between effects of biochemical changes in the central nervous system and changes produced in peripheral nerve fibres. It is known that changes in peripheral nerve occur much later than changes in the metabolism of the c.n.s. From the psychological point of view the effects on behaviour of the latter would seem to be of greater interest than the behaviour of animals whose muscle tonus is abnormal and who show signs of motor inco-ordination due to peripheral causes. The differential effects on behaviour of degree of deprivation were demonstrated in the fourth phase of the present study, water maze learning, where animals evidencing polyneuritic symptoms showed significant impairment in measures of learning when compared with thiamine deficient animals without these symptoms.

The results of previous studies have led to such conclusions as: "It is thus clearly established that deficiency in the vitamin B complex retards the learning process in rats in maze and conditioned-response situations" (Munn, 1950). Our results suggest that these general statements should be amended to include references to the level of deficiency. Alterations in the normal oxidative mechanism of the brain resulting from thiamine deficiency may not affect behaviour until they are sufficiently severe to produce the symptoms of polyneuritis.

Effects of Reduced Caloric Intake

Our results provide no evidence for an impairment of any of the behaviour studied as a result of reduction in caloric intake. The reduction was such as to maintain body weights at approximately the same level as those of animals in the thiamine-deficient group and covered the extended period of somewhat more than 12 months during which the various measures of behaviour were being recorded.

However, there were three instances in which the behaviour of caloric control animals did differ significantly from that of normal control animals fed *ad libitum* on the same diet. The former engaged in a greater amount of locomotion in the open field test, and, in the soluble water discrimination situation, showed more VTE and learned more readily than the latter. The first of these differences is consistent with the well-documented relation between hunger and increased general activity. The second and third suggest that some interaction occurred between the two drives, escape from water and food-hunger, even though the latter was never reinforced at the completion of any trial. Similar observations have been reported by previous investigators (e.g. Elliott, 1929; Muenzinger, 1934; Kendler, 1945; Siegel, 1946; Webb, 1949; Verplanck & Hayes, 1953; Danziger, 1953) interested in the roles played by "irrelevant drives." It has been claimed that the presence of a drive in an animal, required to respond in a manner which does not affect that particular drive but reduces another drive, may influence performance. It may be that the significant effects of the "irrelevant drive" in our study were related to the fact that the caloric control animals, restricted in their food intake, were generally more active and variable in

their behaviour than animals from the other two groups who were fed *ad libitum*. The water discrimination situation as we have used it appears to be one in which greater activity during a trial favours the more rapid breaking of stereotyped behaviour patterns and, hence, more rapid learning of new responses (Russell & Pretty, 1951). Regardless of the validity of this particular explanation, our results show that "irrelevant drives" may have significant effects on performance and indicate the importance of providing adequate controls in designing research in which such effects may be involved.

It might be expected that such interactions between drives would also have been reflected in behaviour in two other situations to which our animals were exposed, the jumping stand and the water maze. Unfortunately, in the former, none of the animals reached the learning criterion within the limit of 300 trials and, therefore, the data cannot be assessed for interaction effects. On the other hand, analysis of performance in the water maze provided no evidence of such effects. Examination of the results of previous investigations indicates that this variation in interaction effects in different experimental situations is not a new observation. In some situations interaction appears to have accelerated learning, in other situations to have had little or no effect, and in still others to have retarded learning. The basis of this variation is not entirely clear, but may be related, first, to the differential effects of different "irrelevant drives" on behaviour and, second, the aspects of behaviour, e.g. perceptual or motor, emphasized in the various experimental situations.

Comparison of Behaviour in the two Insoluble Problem Situations.

Previous research has shown that animals exposed to an insoluble discrimination problem very soon establish stereotyped behaviour patterns which are highly consistent over extended numbers of trials. Figure 3 shows that our animals behaved in this way in the second insoluble problem situation, but not in the first. In the latter the percentage of stereotyped behaviour during any group of ten trials was high, but it did not reach the continued level of 100 per cent consistency characteristic of the former. These differences in behaviour may be reflections of several differences in the two experimental situations. The most prominent difference is in the techniques used in forcing the animals to respond. In the second situation the animals responded in order to escape from water and each response eventually led to escape, to the alleviation of the conditions which initiated it. In the first situation, the animals were forced to respond by being pushed gently from the jumping stand and there was no evidence that they jumped to avoid being pushed. In contrast with the other insoluble problem situation, there was no obvious relation between the result of a response and the circumstances which initiated it, no direct reinforcement of a drive. This suggests that, in the insoluble problem situation, the presence of conditions in which a specific drive is associated with reinforcement favours the establishment of consistent behaviour stereotypes. Our results indicate that thiamine deficiency, prior to the appearance of polyneuritis, had no significant effects on the establishment of such behaviour patterns.

We wish to thank the Medical Research Council for their support of the research. We are also indebted to Mr. A. Summerfield, University College, London, and to Miss A. M. Copping, Queen Elizabeth College, for their advice and assistance.

REFERENCES

- BIEL, W. C. and WICKENS, D. D. (1941). The effect of vitamin B₁ deficiency on the conditioning of eyelid responses in the rat. *J. comp. Psychol.*, **32**, 329-40.

- BILLINGSLEA, F. V. (1942). Intercorrelational analysis of certain behaviour salients in the rat. *J. comp. Psychol.*, **34**, 203-11.
- BRAUN, H. W., RUSSELL, R. W., and PATTON, R. A. (1949). Duration of decrement in learning and retention following electroshock convulsions in the white rat. *J. comp. physiol. Psychol.*, **42**, 87-106.
- DANZIGER, K. (1953). The interaction of hunger and thirst in the rat. *Quart. J. exp. Psychol.*, **5**, 10-21.
- ELLIOTT, M. H. (1929). The effect of appropriateness of reward and of complex incentive on maze performance. *Univ. Calif. Publ. Psychol.*, **4**, 91-98.
- HALL, C. S. (1936). Emotional behaviour in the rat, III. The relationship between emotionality and ambulatory activity. *J. comp. Psychol.*, **22**, 345-52.
- HALL, C. S., and BALLACHEY, E. L. (1932). A study of the rat's behaviour in the field. A contribution to method in comparative psychology. *Univ. Calif. Publ. Psychol.*, **6**, 1-12.
- JAMES, H. (1954). Unpublished research.
- KENDLER, H. H. (1945). Drive interaction: I. Learning as a function of the simultaneous presence of the hunger and thirst drives. *J. exp. Psychol.*, **35**, 96-109.
- KNOPFELMACHER, F. (1953). *The rigidity of position stereotypes in the white rat under varying conditions of reward and punishment*. Doctoral thesis presented to the University of London.
- LAWRENCE, H. D. (1949). Acquired distinctiveness of cues: I. Transfer between discrimination on the basis of familiarity with the stimulus. *J. exp. Psychol.*, **39**, 770-84.
- MAIER, N. R. F. (1949). *Frustration. The study of behaviour without a goal*. New York.
- MAURER, S. (1935). The effect of partial depletion of vitamin B (B_1) upon performance in rats. *J. comp. Psychol.*, **20**, 309-17.
- MOOD, A. MC. (1950). *Introduction to the theory of statistics*. (1950.) New York.
- MUENZINGER, K. F. (1934). Motivation in learning. I. Electric shock for correct responses in the visual discrimination habit. *J. comp. Psychol.*, **17**, 267-77.
- MUENZINGER, K. F. (1938). Vicarious trial and error at a point of choice: I. A general survey of its relation to learning efficiency. *J. genet. Psychol.*, **53**, 75-86.
- MUNN, N. L. (1950). *Handbook of psychological research on the rat*. Boston.
- POE, E., POE, C. F. and MUENZINGER, K. F. (1937). The effect of vitamin deficiency upon the acquisition and retention of the maze habit in the white rat. III. Vitamin B_1 . *J. comp. Psychol.*, **23**, 67-76.
- RUSSELL, R. W. (1953). Behaviour under stress. *Internat. J. Psychoanal.*, **34** (Supplement), 1-12.
- RUSSELL, R. W. and PRETTY, R. G. F. (1951). A study of position habits induced by reward and "frustration." *Quart. J. exp. Psychol.*, **3**, 151-156.
- SIEGEL, P. S. (1946). Alien drive, habit strength and resistance to extinction. *J. comp. Psychol.*, **39**, 307-16.
- STEVENS, H. (1937). Avitaminosis B (B_1), maze performance and certain aspects of brain chemistry. *J. comp. Psychol.*, **24**, 441-58.
- VERPLANCK, W. C. and HAYES, J. R. (1953). Eating and drinking as a function of maintenance schedule. *J. comp. physiol. Psychol.*, **46**, 327-33.
- WEBB, W. B. (1949). The motivational aspect of an irrelevant drive in the behaviour of the white rat. *J. exp. Psychol.*, **39**, 1-14.
- YUDKIN, J. (1949). Studies in biochemical adaptation. An attempt to induce adaptation in rats to high or low intakes of vitamin B_1 . *Internat. Rev. Vit. Res.*, **21**, 245-254.

EXPERIMENTAL DETERMINATION OF LUNEBURG'S CONSTANTS σ AND K

BY

A. ZAJACZKOWSKA

From the Department of Psychology, University College, London

Visual space is, according to Luneburg, a non-Euclidean hyperbolic space of constant negative curvature. Two personal constants enter into the geometrical description of binocular subjective space. One is σ , indicating the degree of depth perception characteristic to the observer. The second is K , the curvature; a value which denotes by how much the visual space of an individual departs from Euclidean space. The constants σ and K were determined experimentally for 30 untrained observers by Luneburg's 3- and 4-point tests. All values found for K were negative and support the hypothesis that the geometry of binocular space is hyperbolic. Some discrepancies between the theory and findings were observed, while the sensitivity of the experiments to the lack of training and even a slight measure of aniseikonia is evident.

I

INTRODUCTION

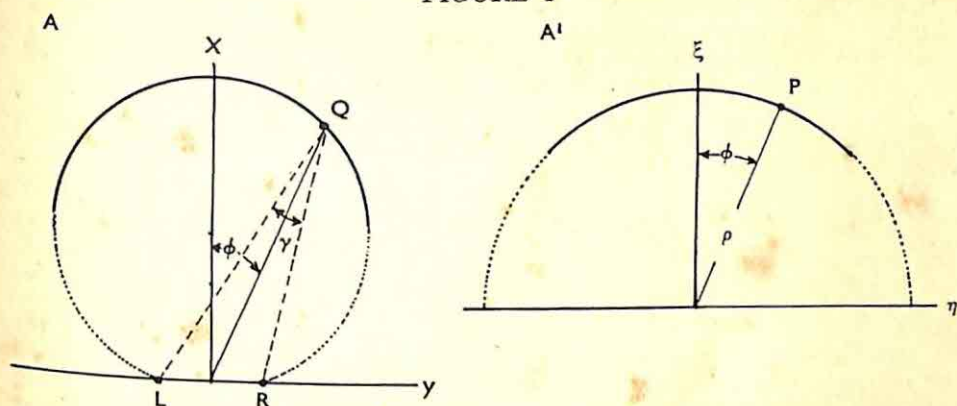
The objective of Luneburg's theory (1947, 1948, 1950) is to establish the relationship between physical space and binocular visual space. The analysis is restricted to the qualities of size, form and localization of objects as they appear to normally sighted observers. In this analysis, firstly, a clear distinction is made between physical space, the space of stimuli, and the visual space, the space of apparent images. Secondly, the analysis refers to spatial relationships in the absence of any indication due to factors such as colour, light and shade, relative size of objects and any qualities of objects recognized from past experience or anticipated owing to the observer's attitudes.

Luneburg concluded that binocular space (in the absence of monocular clues and attitudes), differs from physical space in its metric qualities and is a non-Euclidean hyperbolic space. These conclusions require experimental test. Only two experimental studies have so far been published (Hardy, *et al.*, 1951, 1953), the latter bringing support for the theory. The present study is another attempt to test the validity of the theory.

In Luneburg's analysis of binocular space, for the horizontal plane to which this study is limited, the positions of stimulating points in *physical space* are specified by cartesian (x, y) or by *bipolar co-ordinates* (ϕ, γ). With both systems the y -axis is assumed to pass through the rotation centres of the eyes and is oriented positively to the right. The origin is placed midway between the rotation centres of the eyes. Angle ϕ , the bipolar latitude, can be approximated by $\tan \phi = y/x$ for distances greater than 30 cm. from the eyes. Angle γ , the bipolar parallax (practically equal to the convergence angle) can be expressed by $\gamma = (pd \cos^2 \phi)/x$, where pd is the interpupillary distance. (Figure 1, A).

In *visual space*, the subjective space of an observer, the positions of points will be indicated by cartesian co-ordinates (ξ, η) or by *polar co-ordinates* (ϕ, ρ). The ξ and η axes are assumed to coincide directly with the x and y axes of physical space, and any angle ϕ in visual space is assumed to correspond directly to the ϕ of physical space. The polar co-ordinate ρ (Figure 1, A') refers to the *apparent* distance between the centre of observation, with which the observer identifies himself, and the point considered.

FIGURE 1



The co-ordinates. Bipolar co-ordinates are used for physical space and polar co-ordinates for visual space.

A geometrical description of binocular space, the aim of Luneburg's theory, can be accomplished only in terms of the physical co-ordinates of the stimulating points, as these alone are capable of measurements. Luneburg achieves his goal by postulating a functional relationship between the convergence angle γ and the apparent distance of the stimulating points from the apparent centre of observation. He assumes that the points subtending the same angle γ with the rotation centres of the eyes, that is to say lying on a Vieth-Müller circle (a VMC), appear on the half circle of equidistance of radius ρ (Figure 1). In addition, Luneburg had evidence to suppose that the functional relationship between ρ and γ is exponential and has the form

$$\rho = 2e^{-\sigma\gamma} \quad \dots \quad (1.1)$$

where σ is a constant.

σ is a personal constant varying with individual observers and has been associated with perception of depth. It determines an observer's sensitivity to $\Delta\gamma = (\gamma_1 - \gamma_2)$, i.e. to the increments of radial distance of points from oneself. Irrespective of appreciation of lateral size of objects as expressed by $\Delta\phi = (\phi_1 - \phi_2)$, an observer with good depth perception (large value of σ) will show good judgment of physical distance when estimating line segments extending away from himself.

While the co-ordinates for visual space (ξ, η) or (ϕ, ρ) and formula (1.1) are sufficient to draw a conformal map of visual space, i.e., a map in which the angles are preserved (Figure 1A', Figure 2A', B'), nothing has yet been said as to how distances depicted on such map will appear to an observer in his subjective world. For the calculation of these distances in binocular space, Luneburg's axiomatic system does not limit investigation to only one geometry, but extends over the range of possible experimental answers to a class of Riemannian geometries of constant Gaussian curvature which, besides Euclidean, includes elliptic and hyperbolic geometries. For all these, the shortest distance between two points is measured along a geodesic line which is a straight line only in the case of Euclidean geometry. Thus in the case of hyperbolic geometry the formula giving the distance between any two points in the binocular horizontal plane, the Distance Function, is:

$$\frac{2}{\sqrt{-K}} \sinh \left[\frac{\sqrt{-K}}{2} \frac{D}{C} \right] = \left\{ \frac{\rho_1^2 + \rho_2^2 - 2\rho_1\rho_2 \cos(\phi_1 - \phi_2)}{\left[1 + \frac{K}{4}\rho_1^2 \right] \left[1 + \frac{K}{4}\rho_2^2 \right]} \right\}^{\frac{1}{2}} \quad (1.2)$$

Here D , the distance between any two points of co-ordinates (ϕ_1, ρ_1) and (ϕ_2, ρ_2) is

involved in an expression containing (beside C , a scale factor) the Euclidean distance between these two points given by familiar cosine law and K . The constant K is the curvature of visual space and in the case of hyperbolic geometry is negative. As the value of K approaches zero, D in (1.2) approaches the Euclidean distance between the two apparent points, while for the elliptic geometry (positive K) the same formula remains valid, $\sqrt{-1.K}$ or $i\sqrt{K}$ replacing $\sqrt{-K}$.

K is a personal constant which specifies the shape of geodesic lines along which an observer visually evaluates distances and is in general a value determining his system of ordering visual data depending directly on the two eyes and the visual pathway, as distinct from his knowledge of the stimulating objects acquired by learning processes or anticipated owing to attitudes arising from past experience.

The experimental determination of K is of great theoretical interest. Should the values obtained be found to approximate to zero, the appropriate geometry for the description of binocular space would be Euclidean, other values within the interval $+1$ to -1 would indicate either elliptic or hyperbolic geometry as being appropriate.

The aim of this study has been to obtain a number of experimental determinations of K . The experiments were carried out under dark-room conditions with clues reduced to a minimum, the convergence angle being, in principle, the sole variable.

II

METHOD

The 3- and 4-point tests (Luneburg, 1950, p. 637).

In the *3-point test* three light points, Q_0 , Q_1 and Q_2 are presented. Q_0 and Q_1 are movable on a VMC of parallax γ_0 , Q_2 is movable on a VMC of parallax γ_1 (Fig. 2A). On the map of binocular space (Fig. 2A') apparent points P_0 , P_1 corresponding to points Q_0 , Q_1 , will lie on a circle of radius $\rho = 2e^{-\sigma\gamma_0}$ and point P_2 on a circle of radius $\rho = 2e^{-\sigma\gamma_1}$. The observer is asked to place the point Q_2 in such a position that the distances $D(P_0P_1)$ and $D(P_0P_2)$ appear to him to be equal. The distances Q_0Q_1 are varied and a number of determinations obtained.

If we let

$$X = 4 \sin^2 \frac{1}{2}(\phi_1 - \phi_0) \quad \dots \quad (2.1)$$

$$Y = 4 \sin^2 \frac{1}{2}(\phi_2 - \phi_0) \quad \dots \quad (2.2)$$

$$A = \frac{\rho_0}{\rho_1} \frac{1 + \frac{K}{4} \rho_1^2}{1 + \frac{K}{4} \rho_0^2} \quad \dots \quad (2.3)$$

$$B = \left[\left(\frac{\rho_0}{\rho_1} \right)^{\frac{1}{2}} - \left(\frac{\rho_1}{\rho_0} \right)^{\frac{1}{2}} \right]^2 \quad \dots \quad (2.4)$$

the following simple relation between Y and X results from the Distance Function (1.2):

$$Y = AX - B.$$

Hence a plot of a number of experimental values of Y against X should be linear. The experimental value for B , the intercept with the Y -axis of the straight line best fitting the plot, serves for the computation of σ . It follows from (1.1) and 2.4) that in sufficient approximation

$$\sigma = \frac{\sqrt{B}}{(\gamma_1 - \gamma_0)} \quad \dots \quad (2.5)$$

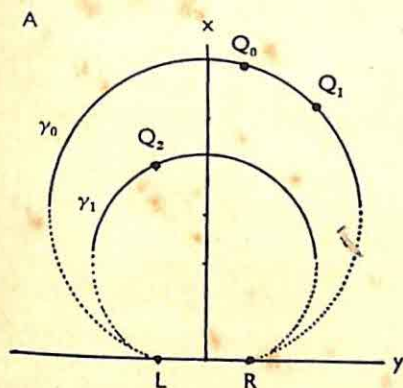
In principle it should be possible to compute the value for K from the expression for A in the 3-point test. The variability of the slope, however, in this experiment is too great to permit K to be determined.

The *4-point test* was designed by Luneburg for determining A and K . In this experiment four points Q_0 , Q_1 , Q_2 and Q_3 are presented. The first pair of these is movable on a VMC of parallax γ_0 , and the second pair on a VMC of parallax γ_1 (Fig. 2B). The observer

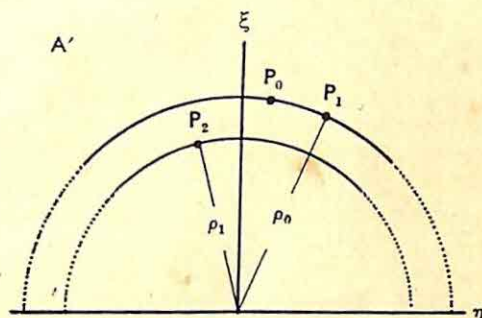
FIGURE 2

THE 3 POINT TEST

Physical Space

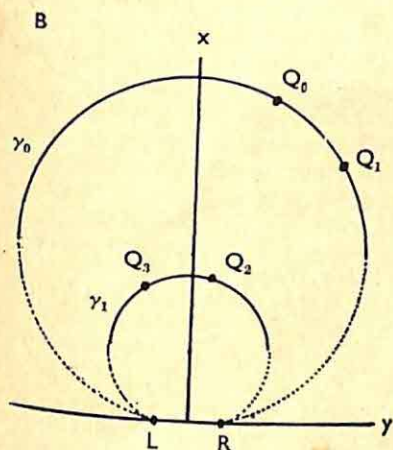


Map of Visual Space

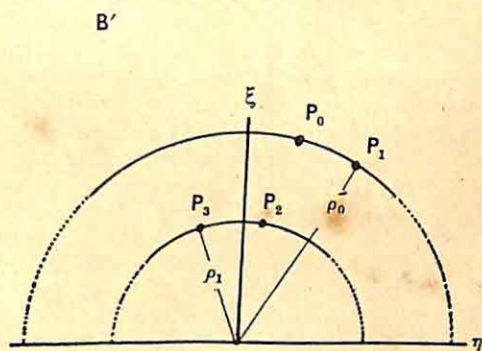


THE 4 POINT TEST

Physical Space



Map of Visual Space



The configurations of stimuli in the 3- and 4-point tests and the corresponding maps of binocular visual space.

can move only the point Q_3 and he is asked to adjust this point so that $D(P_0P_1) = D(P_2P_3)$, the two intervals on the two VMC should appear to him equal.

Letting

$$X = 2 \sin \frac{1}{2} (\phi_1 - \phi_0) \quad \dots \quad (2.6)$$

$$Y = 2 \sin \frac{1}{2} (\phi_2 - \phi_3) \quad \dots \quad (2.7)$$

again from (1.2) we find that here

$$Y = AX$$

where A has the same meaning as in (2.3). The straight line which best fits the experimental values should pass through the origin. With a value for A determined by the 4-point test using (1.1) and (2.3) we obtain

$$K = \frac{2e^{2\sigma\gamma_0} \sinh [\sigma(\gamma_1 - \gamma_0)]}{A - e^{-\sigma(\gamma_1 - \gamma_0)}} - e^{2\sigma\gamma_0} \quad \dots \quad (2.8)$$

The pilot experiment

The success in carrying out the 3- and 4-point experiments depends critically on the experimental conditions adopted. An extensive pilot experiment was carried out before reliable conditions and procedure were established. The results of the pilot experiment can be summarized as follows:

The problems presented to the observer must be geometrically, psychologically and physiologically soluble. The conditions for this are:

(i) The task presented to the observer in the 3-point test is geometrically soluble if for any distance between Q_1 and Q_0 chosen for the experiment the interval between the two VMC is such that

$$\frac{(\rho_0 - \rho_1)^2}{1 + \frac{K}{4} \rho_1^2} < \frac{2\rho_0^2[1 - \cos(\phi_1 - \phi_0)]}{1 + \frac{K}{4} \rho_0^2} \quad \dots \quad (2.9)$$

In other words the interval between the two VMC must not be too large. It must be such that there is a position for Q_2 in which $D(P_0P_2) < D(P_0P_1)$. The experimental conditions would be inadequate if for any position of Q_2 the distance between Q_2 and Q_0 always appears to an observer with good depth perception greater than the distance between Q_1 and Q_0 .

In the 4-point test, on the other hand, it was found that only a considerable interval between the two VMC allows for a reliable determination of A . It follows from (2.8) that for small difference in the diameters of the two VMC such as in case $(\gamma_1 - \gamma_0) \sim 0.01$, $\gamma_0 \sim 0.05$, for anticipated range of constants $+1 > K > -1$ and $0 < \sigma < 20$, the value for A may range only from 1.00 to 1.28. Under these conditions it is evident the experimental error would wholly invalidate the determination of K . No reliable results were obtained for a series of experimental determinations of K carried out with three observers, using $(\gamma_1 - \gamma_0)$ ranging from 0.01 to 0.03, γ_0 from 0.04 to 0.07. In the experimental conditions finally adopted such γ_1 and γ_0 have been chosen as to allow A to range from 1.00 to 3.68.

(ii) It has been found that psychological conditions, providing for behaviour with which Luneburg's theory is concerned are obtained by reducing instruction to the minimum and keeping up a tempo of determinations during the experiment such as not to allow the influx of associations and intellectual processes of the observers, which are not only external to the field on the enquiry but have a pronounced adverse effect on the consistency of measurements. It has been observed that the difficulties encountered in perceiving depth in the 3-point test due for example to the compulsion to make the arrangement of lights symmetric or attempted inferences from the rules of perspective, were not overcome by additional instructions; in such cases frequent reminders to fixate each of the stimuli in turn were effective.

(iii) The conditions physiologically appropriate require avoidance of adverse conditions for fusion which give rise to double images in the area straight in front of the eyes, and limitation of the peripheral range of binocular observation, so that with the head fixed binocular vision remains accurate. While for some adults distinctly linear plots were obtained for stimulus intervals considerably exceeding $(\phi_1 - \phi_0) = 30^\circ$ (about twice as much for the whole stimulus configuration), this range with children had to be much narrower.

The effect of the brightness of lights on the performance of observers has been investigated. It has been found that neither the brightness of lights physically equated, nor a paradoxical arrangement, when nearer lights were dimmer than the farther, affected the results significantly.

Further, the 3-point experiment carried out (a) on various dates under the same conditions, (b) with a varying range of $(\phi_1 - \phi_0)$, (c) using a number of central and peripheral positions ($\phi_0 \geq 0^\circ$), and (d) reversing the order of stimuli ($\phi_2 \geq \phi_0 \geq \phi_1$) confirmed the variability of slope observed by Hardy and his group (1953, pp. 27 and 56), while the intercept of the experimental lines with X-axis tended to remain constant for an observer. Hence the value of slope A calculated on the basis of the 4-point test in conjunction with the X intercept as determined by the least squares method served for the computation of σ .

Finally, the pilot experiment involved an experimental verification of the relationship $\rho = 2e^{-\sigma\gamma}$. Extensive series of the 3-point experiments were carried out with four observers using a range of 5 in one case and 3 different $(\gamma_1 - \gamma_0)$ in the three remaining

cases. For each $(\gamma_1 - \gamma_0)$ 60 determinations were made for six different intervals $(\phi_1 - \phi_0)$. Linear plots of $\log \frac{\rho_1}{\rho_0}$ against $(\gamma_1 - \gamma_0)$ were obtained. This result indicated that formula (2.5) for the computation of σ rests on a satisfactory assumption for use in testing a sample of observers.

Experimental conditions

As a result of the pilot experiments the following experimental arrangements were adopted:

The 3-point experiment was carried out with VMCs of diameters 120 cm. and 100 cm., the average value of $(\gamma_1 - \gamma_0)$ was 0.01. The distances $Q_1 Q_0$ given to be matched by the observer as expressed in value of $X = 4 \sin^2 \frac{1}{2} (\phi_1 - \phi_0)$ were 0.03, 0.05, 0.07, 0.09, 0.11 and 0.13. The values of $(\phi_1 - \phi_0)$ involved ranged from 9.93° to 20.77°. Ten determinations of $Y = 4 \sin^2 \frac{1}{2} (\phi_2 - \phi_0)$ were obtained for each of the six values of X for a central position of the whole stimulus configuration ($\phi_0 = 0^\circ$). Another such series of 60 determinations was carried out with a peripheral position of all stimuli ($\phi_0 = \pm 10^\circ$).

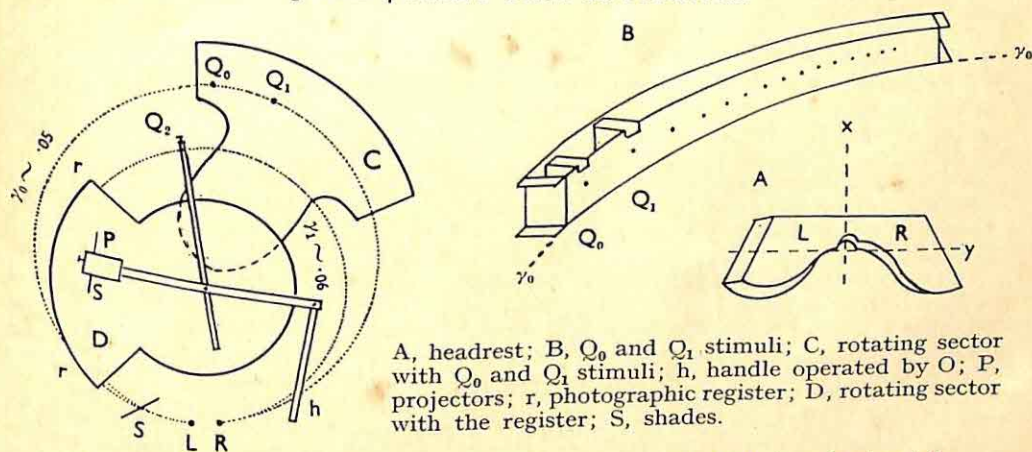
In the 4-point experiment the diameters of the two VMC were 300 cm. and 100 cm., with $(\gamma_1 - \gamma_0) \sim 0.04$. The six distances $Q_1 Q_0$ used when expressed in terms of $X = 2 \sin^2 \frac{1}{2} (\phi_1 - \phi_0)$ were 0.04, 0.06, 0.08, 0.10, 0.12 and 0.14, corresponding to a range of $(\phi_1 - \phi_0)$ from 2.29° to 8.03°, with $\phi_0 = 13^\circ$ and $\phi_2 = 10^\circ$. This arrangement made it possible to utilize the binocular area which on the average was symmetrical to the median for both the $Q_0 Q_1$ as also for $Q_2 Q_3$ intervals. Ten experimental determinations of Y corresponding to each value of X were obtained.

An experimental session consisted of three series each of 60 determinations and lasted about 2 hours during which time the observer remained in the dark room.

Apparatus

The apparatus is illustrated in Figure 3. The observer's face was held in position by a headrest (A) of which the rubber lining was changed to fit individuals. The stimuli used were 6.3-volt lamps shown through uniform 0.8 mm. apertures drilled in aluminium sheet and shaded to diffuse the light. The Q_0 and Q_1 lamps were separated individually and housed at intervals calculated in terms of X in boxes with curved walls (B) fitting an appropriate VMC. The central and the peripheral position of stimuli were exchanged by means of a rotating sector (C) to which the Q_0 and Q_1 lamps were fixed. The Q_2 light was fixed to an arm rotating about the centre of the inner VMC. The handle (h), which the observer used to operate the Q_2 lamp, was attached to the Q_2 indirectly to reduce kinaesthetic cues.

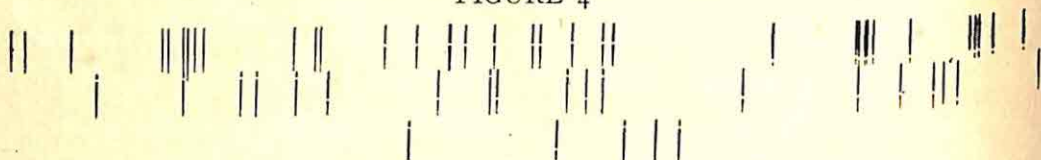
FIGURE 3
3 AND 4 POINT TEST APPARATUS



The 3- and 4-point test apparatus. Diagram on the left is a horizontal projection of the experimental arrangements.

The determinations were registered photographically. By means of an electrical transformation the angular positions of Q_2 and Q_1 light points were recorded simultaneously on a strip of bromide paper fixed to the recording screen (r). A part of an actual register is shown in Figure 4.

FIGURE 4



A part of a photographic register. The black marks indicate the angular position of Q_2 and the white bar the position of Q_1 light point.

An observation screen was mounted above the screen for recording of data; one more projector travelled above the photographic projector (P) in conjunction with the Q_2 lamp. This projected on to a semi-transparent screen a red circle crossed by a vertical bar. On this second screen positions of Q_2 were shown. This device enabled the experimenter to observe the whole course of the experiment, to note whether the instruction had been understood and to interrupt the experiment if, for example, the attention of a child observer was failing.

A rheostat control and separate switches were provided for all lamps. Equal time of exposures for the photographic record was controlled by an electronic timer; this worked in conjunction with electromagnetic counters providing for a definite number (10) of determinations made in a pre-established random order.

Procedure

Thirty observers were tested; they were all children or "naïve" adult observers with three exceptions.

The observer was led to an unfamiliar dark room and the light points were switched on when his head was already in the headrest. Ten minutes for dark adaptation were utilized to familiarize the observer with individual light points and for training. The observer was emphatically reminded several times to use his eyes freely and to look carefully at each light in turn in the course of the experiment.

Thirty determinations in the 4-point test, five for each stimulus interval, were first carried out. Four lights were shown and it was ascertained that all were seen binocularly. Then the Q_1 and Q_0 lights were switched off, and the observer was told that the Q_2 will be called *his light* and that he also *had* one more light which was fixed. Then the Q_1 and the Q_0 were introduced as the *experimenter's* lights, also one movable and one fixed. This was explained by switching the various lights on and off. The brightness of Q_2 was kept as low as possible, consistent with the comfort of the observer. Then all remaining lights were adjusted to the apparent brightness of the Q_2 light. The instruction followed:

"Place *your light* as close to *your fixed light* as *my light* is close to *my fixed light*."

After some training trials, the determinations were registered and counted.

The 3-point test on the central position followed. The Q_0 , now called *the fixed light*, was introduced and the instruction was given:

"Place *your light* as far from the *fixed light* as *my light* is far from it."

If doubt arose as to whether the instruction was understood, more help occasionally was given. This consisted of saying:

"Imagine there is an elastic band between *your light* and the *fixed light*, and that there is another elastic band between the *fixed light* and *my light*. Look carefully at the two elastic bands and make them equal."

Otherwise the help consisted in the repeated request:

"Look with both eyes at *your light*," which was efficacious also in cases of tiredness or undue speed of the determinations. Prolonged hesitation was discouraged. Children carrying out a full series of tests had to be allowed long intervals out of the dark room. After two series of the 3-point test, the remaining part of the 4-point test was carried out.

Computations

The experimental determinations as measured in millimetres on the photographic record were read off in terms of $Y = 4 \sin^2 \frac{1}{2}(\phi_2 - \phi_0)$ or $Y = 2 \sin \frac{1}{2}(\phi_2 - \phi_3)$ from tables prepared for the purpose.

The slope A for the 4-point test was computed using Y as an average of ten determinations for each of six values of X. Under the assumption that the plot should pass through the origin, A was found from formula

$$A = \frac{\Sigma YX}{\Sigma X^2} \quad \dots \quad (2.10)$$

In the 3-point test for each of six values of X an average Y was given by 20 determinations. First, the value of a (the slope of the least squares line through the experimental points) and of b (the intercept of the least squares line with the Y-axis) also $\sigma' = \frac{\sqrt{b}}{(\gamma_1 - \gamma_0)}$ were computed. The value of A, the slope in the 3-point test, was then assessed using the formula (2.8) and the values: σ' (the approximated value of σ), γ_1 and γ_0 of the 3-point test and the slope A obtained in the 4-point test. In practice this operation was carried out again by use of tables worked out for ranges of relevant values. The intercept B for the 3-point test was then given by

$$B = b \frac{A}{a} \quad \dots \quad (2.11)$$

and served in conjunction with formula (2.5) for the computation of σ . K was then given by (2.8).

TABLE I

THE VALUES OF INTERCEPTS, SLOPES AND THE FIRST APPROXIMATION TO σ FOR 28 OBSERVERS

No.	The 3-point test					The 4-point test		
	Slope a	Intercept b	σ'	Slope A	Intercept B	Slope a	Slope A	Intercept b^*
1	0.998	0692	9.98	1.17	0811	2.042	2.151	+ 113
2	1.083	1237	11.12	1.21	1385	2.217	2.975	+ 725
3	0.873	0518	7.44	1.09	0647	1.770	1.455	- 326
4	1.060	1143	12.83	1.22	1315	2.188	2.870	+ 702
5	1.013	0900	10.17	1.13	1004	1.836	1.690	- 050
6	1.160	0840	9.16	1.14	0825	1.385	1.724	+ 325
7	0.842	0374	6.34	1.15	0510	1.837	2.042	+ 210
8	0.945	0408	6.39	1.14	0493	1.494	1.831	+ 322
9	0.974	0790	9.19	1.13	0917	1.605	1.670	+ 066
10	1.240	1431	12.27	1.17	1350	1.982	1.998	+ 017
11	—	—	—	—	—	—	—	—
12	1.040	1688	13.33	1.28	2077	2.716	3.563	+ 870
13	0.446	0622	7.39	1.16	1618	1.739	2.076	+ 348
14	0.842	0498	6.72	1.13	0654	1.525	1.745	+ 220
15	1.172	1691	11.64	1.18	1703	2.074	2.295	+ 227
16	1.005	0366	6.17	1.21	0440	2.817	2.986	+ 173
17	1.021	1170	10.01	1.16	1328	1.620	1.967	+ 357
18	0.934	0688	7.86	1.16	0855	1.922	2.118	+ 202
19	1.050	0558	7.93	1.22	0648	2.392	2.939	+ 563
20	0.850	0443	6.86	1.17	0610	1.901	2.222	+ 335
21	—	—	—	—	—	—	—	—
22	0.802	0146	9.35	—	0000	1.253	1.910	+ 792
23	0.647	0995	0.00	1.14	1832	2.472	2.426	- 047
24	1.441	0993	9.50	1.19	0813	2.176	2.345	+ 194
25	0.937	2028	9.81	1.18	2662	2.748	2.480	- 277
26	0.855	0255	14.73	1.23	0334	1.679	1.874	+ 200
27	0.912	0359	4.59	1.12	0431	1.048	1.477	+ 408
28	1.032	0318	5.83	1.09	0342	1.352	1.658	+ 314
29	0.858	0713	5.20	1.11	0973	2.429	2.112	- 327
30	1.172	0149	7.34	1.17	0134	1.152	1.171	+ 030
			3.52	1.05				

* + above origin; - below origin.

III

RESULTS

The values yielded by the 3- and 4-point tests for a , b , A , B and σ' are listed in Table I for 28 observers. Derivation of constants for observers 11 and 21 were based on more extensive tests. The slopes a and the intercepts b' of the least squares lines for the 4-point tests, calculated without assuming that the lines should pass through the origin, are also given in Table I.

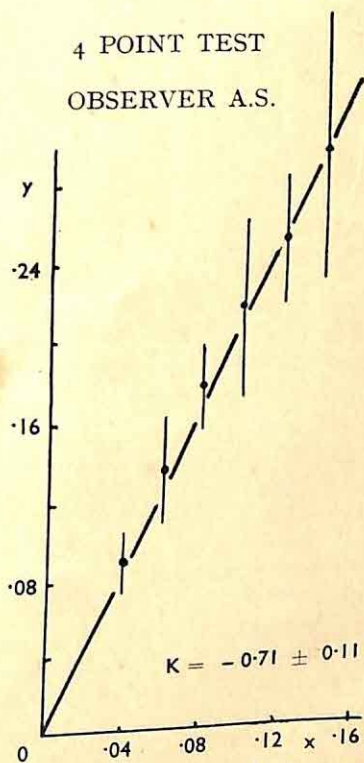
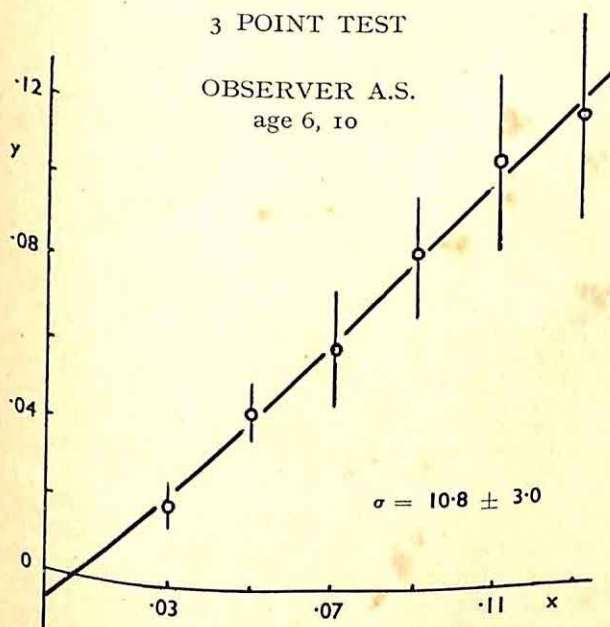
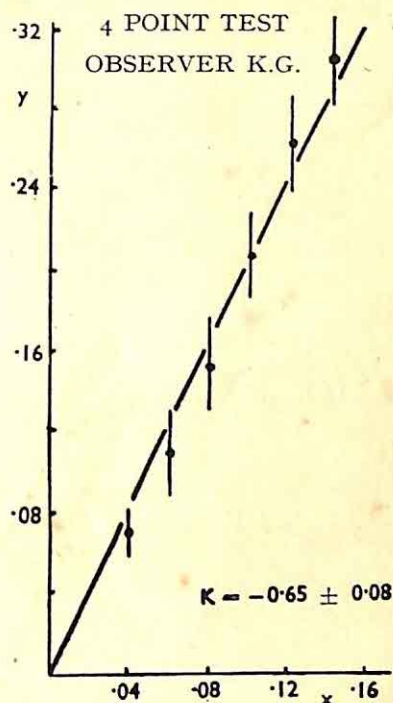
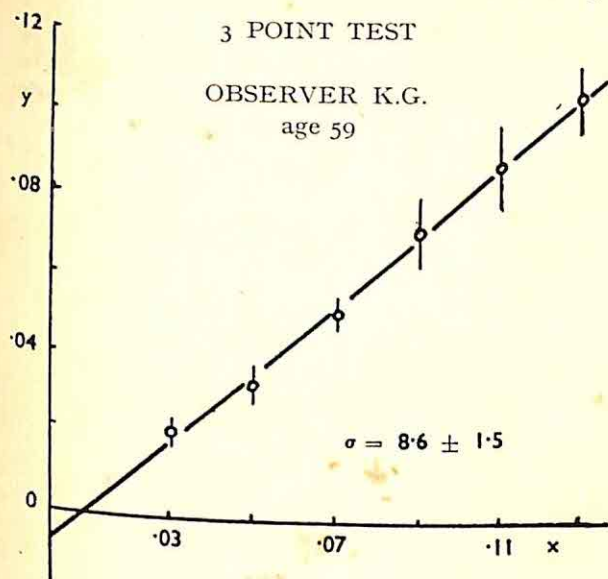
The values of σ and K obtained for 30 observers are given in Table II.

TABLE II
VALUES OF σ AND K OBTAINED BY THE 3- AND 4-POINT TESTS FOR
30 OBSERVERS

No.	Observer	pd (cm.)	Sex	Age (years, months)	σ	K
1	A. S.	5.00	m	6, 10	10.8	-0.71
2	S. H. B.	6.00	f	7, 1	11.8	-0.94
3	M. F.	5.20	m	7, 6	8.3	-0.23
4	A. B.	5.00	f	8, 0	13.8	-0.91
5	M. E.	5.60	f	8, 6	10.7	-0.29
6	J. B.	6.00	m	11, 3	9.1	-0.40
7	S. B.	5.80	f	12, 9	7.4	-0.75
8	G. B.	6.00	f	13, 1	7.0	-0.62
9	B. W.	5.80	m	14, 2	9.9	-0.30
10	J. W.	5.85	f	16, 11	11.9	-0.47
11	J. P.	5.75	m	18, 0	12.4	-0.52
12	E. K.	5.85	f	18, 6	14.8	-1.06
13	B. A.	6.40	m	25	11.9	-0.47
14	S. P.	6.30	m	26	7.8	-0.48
15	T. K.	6.70	m	29	11.7	-0.60
16	M. S.	5.88	f	30	6.8	-0.99
17	A. J.	6.48	m	32	10.7	-0.45
18	M. K.	6.40	f	37	8.7	-0.68
19	S. H.	5.65	f	38	8.5	-0.97
20	M. V. S.	6.25	f	38	7.5	-0.79
21	H. S.	6.30	f	38	9.7	-0.98
22	I. M.	6.30	f	39	0.0	-1.00
23	K. D.	6.30	m	44	12.9	-0.66
24	M. H. B.	6.11	m	44	8.9	-0.79
25	W. K.	5.80	m	47	16.9	-0.55
26	K. K.	6.60	m	48	5.3	-0.72
27	E. F.	6.20	f	48	6.4	-0.33
28	M. B. S.	6.50	m	53	5.4	-0.57
29	K. G.	6.90	m	59	8.6	-0.65
30	S. V. S.	6.60	m	72	3.3	-0.05

The standard errors, expressed as percentages of the observed values are estimated as equal to 20 per cent. for σ and 17 per cent. for K . The error in K is due both to error in the determination of the slope in the 4-point test and to the error in the determination of σ . The standard error in K due to the error in slope in the 4-point test is, however, much smaller and is only about 6 per cent. This last result illustrates well the fact that the 3-point test is much cruder for purposes of estimation than the 4-point test.

FIGURE 5



Plots of Y against X in the 3- and 4-point tests for two observers. Standard deviation marked on the graphs.

The standard deviations of experimental determinations of Y values in the 3- and 4-point tests are indicated for two observers in Figure 5. The values of relative standard deviation were found to be substantially the same for both tests and did not depend significantly on the size of intervals Q_1 Q_0 .

The measurements made on the same interval by subjects in the 3- and 4-point tests were found on the whole to be normally distributed.

No significant difference was found with respect to slope a in the 3-point test for the central ($\phi_0 = 0^\circ$) and the peripheral ($\phi_0 = \pm 10^\circ$) position of the stimulus configurations.

A trend was observed for depth perception to deteriorate with age (in particular in the group of female observers) while the absolute values of K were found to increase insignificantly with age. The difference between the mean value of K for female observers (-0.77) and the mean value of K for male observers (-0.51) is highly significant. The seven highest values of K were obtained for female observers.

The important result of this investigation is that the values found for K , the curvature of visual binocular space, range from -1.06 to -0.05 supporting the hypothesis that the geometry of binocular space is hyperbolic.

IV

DISCUSSION

Two discrepancies between the theory and the results obtained will be noted:

(i) The slopes of the least squares lines in the 3-point test were in general smaller than the theoretical slopes, the difference between the mean slope obtained ($a = 0.987$) and calculated on the basis of the 4-point test ($A = 1.163$) is highly significant (0.01 level).

(ii) The line passing through experimental points in the 4-point test should according to the theory pass through the origin. It is estimated, however, that the mainly positive intercept with Y -axis (b' in Table I) is in half the cases significant.

While noting these discrepancies with the theory it is to be observed, that the use of the 4-point test slope A in place of the 3-point test's a for the determination of σ did not affect the value of σ considerably. As seen from Tables I and II, the approximated values of σ are generally somewhat smaller, thus the values of K corresponding to them would differ even more from zero than the values of K as given in Table II. On the other hand this last effect has been counteracted by the second irregularity observed; namely $A > a$ in the 4-point test. The values of K worked out for the same sample of observers using only the least squares values b , σ' and a from the 4-point test, with few exceptions hardly differed from those given in Table II.

It is to be noted further that in the cases of several observers who took part in a number of experimental sessions during the pilot experiment, the acquired training resulted in:

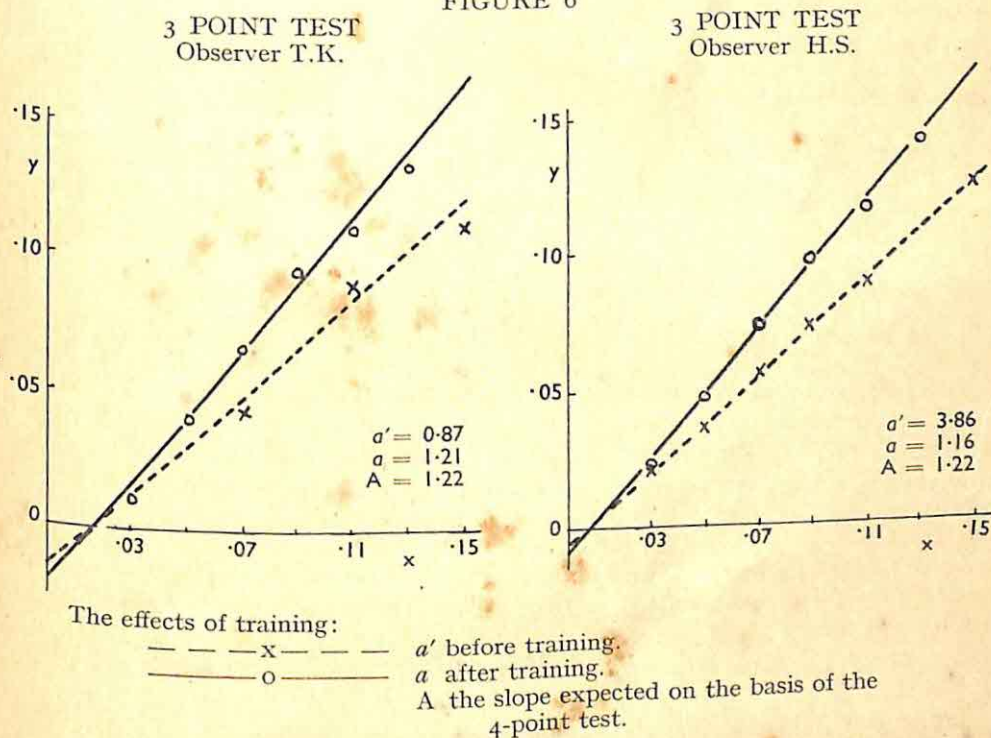
(i) Almost complete disappearance of relevant discrepancy between the slope obtained by the 3-point test as compared with the slope expected on the basis of the 4-point test (Figure 6).

(ii) A wider range of intervals ($\phi_1 - \phi_0$) for which the plot was linear in the 3- and 4-point tests.

Experiments carried out repeatedly with some observers over a period of two years indicated that the use of the 4-point test only for the determination of slope for both experiments tended to eliminate factors such as incidental tiredness, the lack of training, asymmetry of the eyes, even defective experimental conditions, which particularly affect the 3-point test.

The effect of the common asymmetry between the eyes on the performance in the 3-point test is very pronounced in particular in the case of poor depth perception. Luneburg's hypothesis is based on the assumption that the size of frontal images on the two retinae is the same. Differences in size between two retinal images are not infrequent; the percentage difference of magnification is approximately 1.4 per cent. per diopter of anisometropia and may range up to 2 per cent. per diopter of correction in cases of astigmatism (Ogle, 1950, p. 264). The effect, however, of an afocal magnification lens, for instance on the slope in the 3-point test, can be predicted on the basis of Luneburg's theory. Such predictions were worked out, indicated much more pronounced distortion in stereoscopic perception of depth for $K = -1$ than for $K = 0$, and were to a large extent confirmed by actual experiments on two observers. It has been predicted, for example, that for $\sigma = 10$, $K = -1$, the least squares slope in the 3-point test (under experimental conditions as used for the sample for 30 observers) with a 5 per cent. magnification size lens in front of the right eye should be $a = 3.595$.

FIGURE 6



For the same conditions but with the same lens in front of the left eye the predicted slope is $a = 0.703$ ($b = -0.0138$). The actual experimental results for observer H. S., $\sigma = 9.7$, $K = -0.98$ were: with 5 per cent. magnification size lens in front of the right eye $a = 4.092$, with the same lens in front of the left eye $a = 0.788$ ($b = -0.0138$). This last result compares well with the result obtained in the same test carried out under normal conditions with observer I. M., who has aniseikonia, the images on her left retina being larger. The slope found for her was $a = 0.802$, and as in case of observer H. S. with size lens in front of the left eye, the value b was negative ($b = -0.0146$). In general, in cases of poor depth perception the value for b is small and will be affected even by a slight measure of aniseikonia, particularly in cases of large absolute values of negative K .

To summarize our discussion, while the control of the experimental conditions so as to obtain binocular reactions based on the differences in the convergence angles is a highly complex task, we also encounter the difficulty that the observers do not have training in such behaviour and have formed a habit to rely on non-binocular factors; further, even a slight asymmetry between the eyes, a very common occurrence, will affect the results.

Considering these difficulties and that with regard to binocular behaviour $K < -1$ is impossible (such a value would imply that the apparent size of a receding object could first increase with distance and then become infinite and negative) it is remarkable that with one exception no experimental values for K were obtained outside the interval from 0 to -1 .

Further, the plots of Y against X obtained in 3- and 4-point experiments were on the whole distinctly linear. Some deviations in this respect were noted in the case of children, but these were not found to depart significantly from linearity. The linearity of the plots, as shown by Luneburg (1950, p. 639) is an experimental verification of the hypothesis of constant curvature.

Lastly, a number out of our 30 observers took part in the Horopter and the Alley experiments (Zajaczkowska, 1955) and it was found that within certain limits of confidence the values of personal constants σ and K reported in this paper are not devoid of predictive value.

In conclusion, while discrepancies between the theory and findings were noted the experiments performed favour the hypothesis of negative curvature of binocular space.

This paper reports part of a more general study of Luneburg's theory carried out for the Degree of Ph.D. in Psychology of the University of London.

This research has been carried out under grant from the Central Research Fund of the University of London. I would like to thank most gratefully Mr. A. R. Jonckheere for his very generous assistance at every stage of the investigation, Mr. J. W. Whitfield for valuable criticism, Mr. W. Kunowski for solving the electronic problems of the apparatus and much helpful discussion on the mathematical structure of the theory, Mr. R. J. Audley for his comments on the manuscript, Captain M. H. G. Billson for reading through the text, and the observers for their conscientious effort.

REFERENCES

- LUNEBURG, R. K. (1947). *Mathematical Analysis of Binocular Vision*. Princeton, N.J.
 LUNEBURG, R. K. (1948). *Metric Methods in Binocular Visual Perception, Studies and Essays*. Courant Anniversary Volume, pp. 215-240. New York.
 LUNEBURG, R. K. (1950). The metric of binocular visual space. *J. Opt. Soc. Amer.*, **40**, 627-642.
 HARDY, L. H., RAND, G., and RITTNER, M. C. (1951). Investigation of visual space. The Blumenfeld alleys. *Arch. Ophthalmol.* (Chicago), **45**, 53-63.
 HARDY, L. H., RAND, G., RITTNER, M. C., and BLANK, A. A. (1953). *The Geometry of Binocular space Perception*. The Knapp Memorial Laboratories, Columbia University, New York.
 BLANK, A. A. (1953). The Luneburg theory of binocular visual space. *J. Opt. Soc. Amer.*, **43**, 717-727.
 OGLE, K. N. (1950). *Researches in Binocular Vision*. Philadelphia.
 ZAJACZKOWSKA, A. (1955). An experimental study of Luneburg's theory of visual perception. Univ. of London, Thesis. (Unpublished.)

"HANDEDNESS" AND SKILL

BY

K. A. PROVINS

From the Unit for Research on Climate and Working Efficiency of the Medical Research Council, Department of Anatomy, University of Oxford.

The variability of handedness with different tasks is discussed. Experiments are described which show under what conditions handedness becomes evident. Tasks involving three different levels of complexity were used. The simplest task measured the accuracy with which a particular pressure could be reproduced in isometric contraction of the flexors of the index finger on each side in 21 female subjects. In the second situation, the maximum speed of making an attempted tapping movement under the same conditions, was measured in ten of the same subjects using the same muscle group alternating with its antagonists. The same ten subjects were also tested on an aiming task which provided the third level of complexity. The results suggest that differences in performance between the two sides only occur where "timing" or the serial organization of muscle activity is required and that such differences may be due to training.

Whether handedness is inherited or acquired is briefly discussed, and a second series of experiments using the same tasks as before were carried out on one female and nine male subjects. In this instance, the first two tests were used on the big toe of each side as well as the index finger. The results confirm that differences in performance between the two sides on these tasks can be adequately explained in terms of usage or training.

The hypothesis that "timing" is therefore important in the learning of any movement where serial muscle contractions are involved was tested and confirmed in a third experimental series. The consistency of timing of the application of force in turning a crank handle at maximum speed was measured in five male subjects before and after training. The implications of the results are discussed in relation to other researches on skills.

INTRODUCTION

The importance of unilateral dominance for neurological diagnosis has been stressed by Brain (1945). Similarly, the implications for education have recently been discussed by Hildreth (1949 a, b, c., 1950 a, b). A relatively new aspect of the problem, however, has become prominent with the advent of large and complicated industrial machines which require considerable care in the design and layout of controls for optimum human operation. Where these controls are numerous or have to be positioned for operation by either one hand or the other, the question naturally arises whether they may be operated better by the right or left hand. If there is any choice in the matter, the control is usually placed for operation by the right hand, with the result that left-handed operators may be handicapped.

Differences in performance between the preferred and non-preferred hand vary considerably from one task to another (Jack, 1895; Durost, 1934; Blau, 1946; Burge, 1952), but little is known of the reason for this. Several investigators have suggested that the degree of lateral dominance displayed in a specific situation is dependent on the complexity of the task (Jack, 1895; Durost, 1934; Blau, 1946), but little indication is given of what is implied by complexity. The typical tracking situation has been considered to require a high degree of skill (Gibbs, 1954), yet one hand has been shown to perform this task as well as the other (Grant and Kaestner, 1955).

The present series of experiments was therefore designed to determine rather more closely the conditions under which handedness is manifest. In the first instance the

type of movement tested was made as elementary as possible by confining it to attempted flexion of the index finger at the metacarpo-phalangeal joint. Actual displacement of the finger was virtually eliminated by measuring the accuracy with which a particular pressure could be reproduced against a spring-steel bar. A previous study using this technique (Provins, 1955) provided evidence that, unlike slower rates, the fastest rate of application of finger pressure is not subject to modification by sensory control, so that a comparison of the two sides at each rate would indicate whether handedness was due to differences in sensory or motor aspects of the task. For example, any differences in the quantity or quality of the sensation derived from the two sides would presumably be evident in differences of performance at the slow but not at the maximum rate of application. On the other hand, differences in the ability to grade the muscle contractions on each side would be evident at both rates of application. It was realized, of course, that if differences were evident between the two sides at both rates of application, it would be extremely difficult to determine if they were due to afferent and efferent differences or only to differences concerning the motor aspect of the two sides.

Performance on each side at two other tasks was also tested. The first of these recorded the maximum rate of attempted oscillation of the index finger using the same apparatus as in the first experiment, except that the finger was strapped to the spring-steel bar with cello tape so that the actual movement of the finger was again negligible. In the third experiment, the accuracy of aim was recorded in throwing darts at a standard dartboard in the usual manner. This constitutes the first experimental series.

The second series of experiments was undertaken to confirm on a wider basis the findings of the initial experiments which suggested that differences between the two sides only occur where timing of the component parts of a movement become important, and that such differences may be due to unequal amounts of training. The tests used were essentially the same as in the first series.

In the third series of experiments, the patterns of applied force in turning a crank handle with the same hand before and after training were analysed to test the hypothesis that training improves the consistency of "timing" of the appropriate muscle contractions in a task demanding serial organization of muscle activity.

APPARATUS

The experimental apparatus used in the experiments concerning accuracy of applied pressure and speed of tapping comprised, basically, a spring-steel bar, on each side of which was mounted a 1,000 ohms strain-gauge element. These two strain gauges formed two of the resistances of a Wheatstone bridge which was supplied by a 2 Kcs. oscillator; the output from the bridge was led through an A.C. amplifier and phase-sensitive rectifier before being amplified further by a D.C. amplifier and recorded on a Brush pen recorder. The natural frequency of the system was 21 cycles per second.

One end of the spring-steel bar was held rigid in an adjustable frame which allowed the apparatus to be manipulated to the required position for each subject. To the other end of the steel bar was attached a brass disc of particular dimension, against which the subject applied the pressure required. The size of the brass disc in each instance is specified in the experiments described later.

At the beginning of each experimental session, the apparatus was calibrated by turning the spring-steel bar horizontal and placing in a pan fitted to its free end, a weight equivalent to the deflecting force used in the subsequent experiment. The recorder was then adjusted to give 25 mms. deflection of the pen. The actual forces used are given in the description of the individual experiments.

In the aiming task described, an ordinary standard dartboard was used, the centre being 5 ft. 6 in. above the floor and the subject standing 8 ft. 6 in. distant from it.

SERIES I

METHODS

Experiment 1. The accuracy with which a particular pressure could be reproduced with the index finger of each side was compared at both the maximum and a slower rate of application. The gradual rate of application was in the subject's own time, which was usually of the order of 2-3 seconds duration.

Twenty-one student occupational therapists, all females, were used as subjects. They were seated in a comfortable posture with the trunk upright and the lower arm and ulnar border of the hand being tested supported on the padded chair arm. Pressure was applied by the palmar aspect of the distal phalanx of the index finger when making an attempted movement at the metacarpo-phalangeal joint by contracting the finger flexor muscles. In producing a pen deflection of 25 mms., the brass disc of radius 0.25 in. was displaced 0.006 in. so that the muscle contractions involved may be considered isometric. The apparatus was calibrated to produce 1 mm. pen deflection per gram weight applied to the brass disc.

The subject was told that whenever she exerted a pressure against the brass disc, the recording pen would be deflected in direct proportion to the force applied and that the force she was required to apply in every instance corresponded to a 25-mm. pen deflection. The subject was unable to see the recording pen and was told that the first time that she applied pressure to the brass disc she should do so gradually and carefully, and the pen deflection so produced would be called out to her as she progressed every five mms. In all subsequent runs she would have to rely entirely on her own judgment and at the slow rate of application, apply a continually increasing force in her own time until she considered that she had just reached 25 mm., then to relax and take her finger away from the disc surface.

Instructions concerning the "fast" run were merely that the application should be at the maximum rate possible, and care was taken to examine the form of the deflection after each contraction, only sharply defined single spikes being accepted.

On being given the warning "stand by," the subject brought her finger into position so that the skin area specified was not quite touching the brass disc. When she was told "now," she closed her eyes to eliminate the possibility of using visual aid, and started applying the required pressure. After she had completed the contraction and returned to the relaxed position, she was told the deflection achieved to the nearest 0.5 mm. Fifteen seconds interval was allowed between the start of one "trial" or "contraction" and the start of the next. Ten practice trials were given before recording each run, which consisted of a further ten contractions. The actual pen deflection achieved in each instance was noted to enable the mean error for the ten contractions to be calculated.

The order in which the right and left hands and the fast and slow rates of application were used was randomized to reduce to a minimum the possible effects of practice.

Experiment 2. The maximum speed with which ten of the same subjects in the same posture could oscillate the index finger in attempted flexion and extension at the metacarpo-phalangeal joint was recorded on each side. On this occasion, the forearm was strapped firmly to the arm of the chair, the hand gripped a rigid vertical handle and the distal phalanx of the index finger was strapped tightly to the brass disc.

The actual amount of displacement of the brass disc for the maximum force required was 0.096 in., so that the muscle contractions involved may again be considered as isometric.

The subject was informed that maximum speed rather than accuracy was the aim and that she would be told if the amplitude of her tapping was very different from the 25 mms. pen deflection required. Accordingly, during the actual run which usually lasted about 15 seconds, the experimenter occasionally corrected the subject if the average amplitude was below about 10 mms. or above 35 mms. peak to peak.

Three different levels of applied pressure were used so that in each instance 25 mms. pen deflection corresponded to forces of 25 gms. weight; 100 gms. weight and 400 gms. weight. The order in which these were tested on each side was randomized, half the subjects being tested on the right side first and the other half on the left.

Experiment 3. The subjects used in experiment 2 were also tested on both sides for the accuracy with which they could perform in an aiming task using a standard dartboard in the usual manner. In this instance, subjects were requested to aim for the centre of the board every time and to throw darts singly, the error of aim being measured and recorded for each dart, which was withdrawn before the next was thrown 15 seconds later.

Half the subjects were tested first with the right hand and half were first tested with the left.

RESULTS

Experiment 1. The differences between the mean errors (in mms.) of 21 subjects for the two sides at both the gradual and maximum rates of application were tested for significance by "t" test. The mean values are given in Table I, together with the results of the "t" tests.

TABLE I

COMPARISON OF THE MEAN ERRORS (IN MMS.) OF THE PREFERRED AND NON-PREFERRED HAND IN 21 SUBJECTS IN REPRODUCING A PARTICULAR PRESSURE AT TWO RATES OF APPLICATION

<i>Fast Application</i>		<i>Slow Application</i>	
<i>Preferred</i>	<i>Non-Preferred</i>	<i>Preferred</i>	<i>Non-Preferred</i>
12.69	12.70	6.25	6.83
Not significant $t = 0.007$		Not significant $t = 0.91$	

Although the mean errors at both the fast and slow rates of application were less when the preferred hand was used, the differences are not significant.

Experiment 2. The oscillograph records of the finger-tapping were analysed over a period of ten seconds and the rates derived in cycles per second. The differences in rate between the two sides at the three different force values for the ten subjects were tested by "t" test. The results are given in Table II together with the mean values in cycles per second.

TABLE II

COMPARISON OF THE MEAN RATE (IN CYCLES PER SECOND) OF FINGER OSCILLATION ON THE TWO SIDES IN 10 SUBJECTS AT THREE DIFFERENT SENSITIVITIES

25 gms.		100 gms.		400 gms.	
<i>Preferred</i>	<i>Non-Preferred</i>	<i>Preferred</i>	<i>Non-Preferred</i>	<i>Preferred</i>	<i>Non-Preferred</i>
5.045	3.806	4.735	4.042	4.324	3.818
Significant $P < 0.001$		Significant $P < 0.02$		Significant $P < 0.01$	

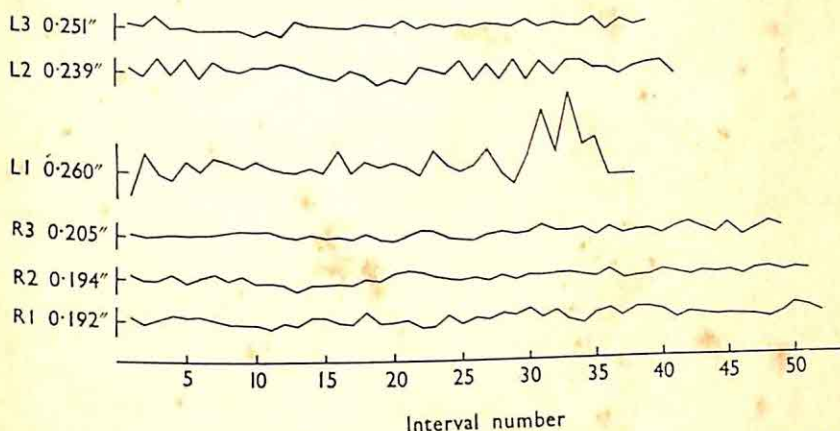
There is a significant difference in speed of tapping between the two sides at each force value tested in favour of the preferred hand.

The consistency of cycle length on the two sides was determined by measuring the inter-peak distance for successive cycles throughout the ten-second period previously analysed on all records and plotting the results as in Figure 1.

The values plotted are for one subject only, but the trend is considered representative of the results as a whole. There is very little difference in consistency of cycle length at the different force values with the preferred hand, although the speed of tapping is greatest at the most sensitive calibration. With the non-preferred hand, however, the consistency of cycle length varies appreciably from one force value to another and is best at the least sensitive calibration. Variation in cycle length is most marked in the non-preferred hand at the most sensitive calibration which also corresponds to the lowest tapping rate.

Experiment 3. The mean errors, in cms., are given in Table III together with the result of the "t" test of the differences between the two sides for the ten subjects.

FIGURE 1



Variation of tapping interval on the two sides for one subject tapping at maximum speed with the index finger at three different sensitivities for ten seconds. The mean tapping interval (in seconds) is given in each instance. L = left side; R = right side. In conditions 1, 2 and 3, calibration values = 25 gms., 100 gms., and 400 gms. weight respectively for 25 mm. pen deflection.

The mean error was less for the preferred hand, but the difference between hands was significant at only the 10 per cent. level.

TABLE III

COMPARISON OF THE MEAN ERRORS (IN CMS.) OF THE PREFERRED AND NON-PREFERRED HAND IN 10 SUBJECTS IN AN AIMING TASK

<i>Preferred</i>	<i>Non-Preferred</i>
12.29	15.85
Significant $P < 0.1$	
$t = 2.06$	

When questioned on completion of the experiments 20 of the 21 subjects in the first experiment said that they were right-handed and one left-handed, although none expressed a particular preference in this task. In the second experiment, hand preferences corresponded in each instance with a better mean performance, but in the aiming task this was not always so. In fact three subjects showed a better average with the non-preferred hand, although only in one subject was this appreciable.

DISCUSSION

In these three experiments, the different tasks may be considered as providing three different levels of motor complexity. Apart from the fixing activity of a large number of muscles in all three tasks, in the first experiment the movement was probably produced by little more than the graded contractions of the finger flexors. In the second and third experiments, however, the required result could only be achieved by the contraction of several muscles in a particular order. In the second

experiment this serial muscle activity involved the alternate contraction of the finger flexors and extensors with the interval between the contractions of antagonistic muscle groups as short as possible. In the third experiment, many more muscles were used not only in a particular order, but probably with critical time intervals between the contractions.

It is true that in the first two experiments the amount of actual movement of the finger was negligible, while in the aiming task quite appreciable displacement of the limb occurred. It must also be noted that in the first and third experiments the performance was scored in terms of the accuracy of execution of the movement, while in the second task speed of movement was the criterion. Yet neither of these facts can account for the results which vary independently of these conditions. On the other hand, the records of consistency of cycle length in the speed of tapping task do support the view that differences in the serial organization of the muscle activity required are responsible for the differences between hands where they occur. Certainly there is a close correspondence between speed of tapping and consistency of cycle length so that the more consistent the time interval between taps the higher the rate of tapping.

The results suggest therefore, that in making simple muscle contractions, one hand is no better than the other providing the factor of "timing" is not present. In any movement requiring the serial organization of muscle activity involving, that is to say, the timing of the component parts of the movement, then the hand used becomes an important factor. Thus, in experiment I, in either the fast or slow runs, there was no detectable difference in accuracy between the two sides, while in the tapping series where the contractions of the agonist and antagonist muscles had to be carefully timed to alternate with one another, the difference was highly significant. Similarly, in the aiming task, many muscles are used serially, and it seems likely that the same factor of timing of the right moment for each muscle to act tended to disclose a difference between hands.

Yet it is possible that handedness, implying some form of inherited or inborn lateral dominance, was not the determining factor at all. The differences could more easily be attributed to usage or training which would account for the doubtful significance of the differences between sides in the aiming task, in which girls do, on the whole, receive little practice. Certainly all of the present subjects except one said they had rarely, if ever, played darts before. The exception was one of the only two subjects who showed a marked superiority of the preferred hand in this task.

It has previously been suggested by several authors that handedness is the result of the environment and not heredity. To test this hypothesis in the tasks already used a second series of experiments was carried out.

SERIES 2

INTRODUCTION

The strongest claim for a genetic basis for handedness comes from the study of dermatoglyphics, for, as Newman (1934) points out: "After an extensive search for morphological criteria that might be useful indicators of inborn differences between functional right- and left-handers, we have been driven by a process of exclusion to use palm, finger and sole papillary patterns as the only available material for this study. . . ." "The main advantage for this study of palm, finger and sole patterns lies in the fact that they are determined and fixed for life during the third month of foetal life and are not at all influenced by post-natal functioning of the hands." Consequently, the demonstration that there is an association between the asymmetrical distribution

of palm and finger patterns and functional handedness (Cromwell and Rife, 1942; Rife, 1943) seems to provide conclusive evidence that at least some part of handedness is inherited (Rife, 1950). However, it has recently been shown by Holt (1954) that although the ridge count characteristic of finger prints is inherited, its asymmetrical distribution between the fingers of the right and left hands is not inherited, which must cast serious doubt on the validity of the assumptions underlying previous work relating dermatoglyphics to handedness.

Evidence from other fields on this question is again inconclusive, but suggestive that training at least plays the major role in determining handedness. Hildreth (1949c), in reviewing the literature on handedness with particular reference to child development concludes, "The most that can be said is that heredity furnishes the normal human infant with two hands and arms that are about equally responsive to training, and an active mind that is capable of learning complex motor adjustments. The explanation that best seems to fit the facts is that handedness is the product of developmental integration. People are right- or left-handed because they have learned to be, not because they were born that way."

The evidence from animal studies seems to indicate that in those species which display one-sided dominance, there is no consistent preference for one side or the other (Cole, 1955), while observations from neurological material tend to show in ever-increasing numbers that one-sided cerebral dominance is an over-simplification of the facts (Glees and Cole, 1952; Goodglass and Quadfasel, 1954; Zangwill, 1954). It has further been suggested by Blau (1946) that such cerebral dominance is the result rather than the cause of handedness.

Consequently, the second series of experiments was designed to determine whether differences in performance on the two sides in the three tasks previously described could be accounted for by training. Subjects were tested in the accuracy of reproduction of pressure using two different body members—the index finger and the big toe. It was considered that while the index finger was normally offered ample opportunity for differential training on the two sides, as the feet were usually completely enclosed in some form of footwear, any differences in performance here could hardly be explicable on the same basis. The use of practically all male subjects in this series was adopted mainly to confirm or reject the conclusions of the first series concerning the doubtful significance of the difference in performance of the two sides in the aiming task. However, it must be stressed that these subjects were not selected for their dart-playing ability, which was quite unknown to the author prior to testing.

METHOD

Nine males and one female, all members of the Department staff, were used as subjects throughout. All subjects said that they were right-handed.

Experiment 1. The accuracy of application of pressure on the two sides was compared at a gradual rate of application using the big toe. Subjects made each contraction in their own time, which was usually of the order of 2-3 seconds duration.

Subjects lay in a supine posture on a well-padded table-top with the plantar surface of the bare foot being tested strapped firmly against a flat wooden support. The big toe from the metatarso-phalangeal joint projected beyond the end of the board and was quite free to move.

The strain-gauge apparatus used in the first series of experiments was adjusted so that the plantar aspect of the distal phalanx of the great toe was just clear of the brass disc (radius 0.5 in.). Subjects were required to attempt flexion of the big toe at the metatarso-phalangeal joint. The apparatus was calibrated to give 25 mm. pen deflection for 100 gms. weight applied for all subjects, except one who experienced great difficulty in grading the applied pressure, so that he used the less sensitive calibration of 25 mms. for 400 gms. weight applied. The actual amount of displacement of the brass disc for the two calibrations was 0.024 in. and 0.096 in. respectively, so that the contractions can be considered as isometric. The

procedure was as previously described for the finger, except that subjects were allowed to keep their eyes open all the time, as it was considered highly improbable that they could gain any visual indication of pressure applied in this posture.

Half the subjects were tested first on the right side and half were tested first on the left.

Experiment 2. The same apparatus and posture was used as in experiment 1, but this time the speed with which subjects were able to tap the big toe on the right and left sides was compared. The same calibration was used as for experiment one and the same method as that given in experiment 2, first series, for the finger-tapping experiment, except that in this instance the toe was not strapped to the brass disc. This was considered advisable in case the movement became too erratic and the apparatus consequently damaged.

Experiment 3. The accuracy of application of pressure on the two sides was compared using the maximum rate of application with the index finger. The details of the experiment were otherwise as in experiment 1, first series. Half the subjects were tested first on the right side and half were first tested on the left.

Experiment 4. The speed with which the ten subjects could tap with the index finger on each side was recorded in the same way as for experiment 2 in the first series, except that only one calibration value was used. The apparatus was adjusted so that 25 mms. pen deflection corresponded to an applied force of 25 gms. weight.

Experiment 5. Finally, the ten subjects were tested on the aiming task described in experiment 3 of the first series. The method was precisely the same as before.

RESULTS

The mean values of the two sides for the ten subjects in each experiment are given in Tables IV, V and VI, together with the results of the appropriate "t" tests of significance.

Comparing the performance of the right and left big toes, there is no significant difference between the sides irrespective of whether accuracy of reproduction of pressure is compared or speed of a reciprocal movement as can be seen in Table IV. In fact, it is perhaps surprising to note that the mean error of reproducing a particular pressure is less with the left big toe (which is true of six of the ten subjects used), while the mean tapping rate is very slightly in favour of the right side.

TABLE IV
COMPARISON OF THE MEAN ERRORS (IN MMS.) OF THE RIGHT AND LEFT SIDES IN 10 SUBJECTS IN THE REPRODUCTION OF A PARTICULAR PRESSURE AND MAXIMUM RATE OF TAPPING (IN CYCLES PER SECOND) WITH THE BIG TOE

Pressure (slow application)		Tapping	
Preferred (right)	Non-Preferred (left)	Preferred (right)	Non-Preferred (left)
8.84	7.37	2.75	2.73
Not significant $t = 1.743$		Not significant $t = 0.207$	

There is no significant difference between the two sides in the accuracy of reproduction of finger pressure as can be seen in Table V, although the mean error is less on the preferred or right side. Table V also shows that the mean rate of tapping with the index finger of the preferred hand is significantly higher than with the left, in contradistinction to toe-tapping.

The oscillograph records of the tapping experiments were analysed over a period of 10 seconds, and the rates derived in cycles per second as before. The difference in consistency of cycle length between the two hands confirmed the results of experiment 2 in the first series. Differences between the big toes on each side, however, were difficult to discern and sometimes favoured the left side and sometimes the right.

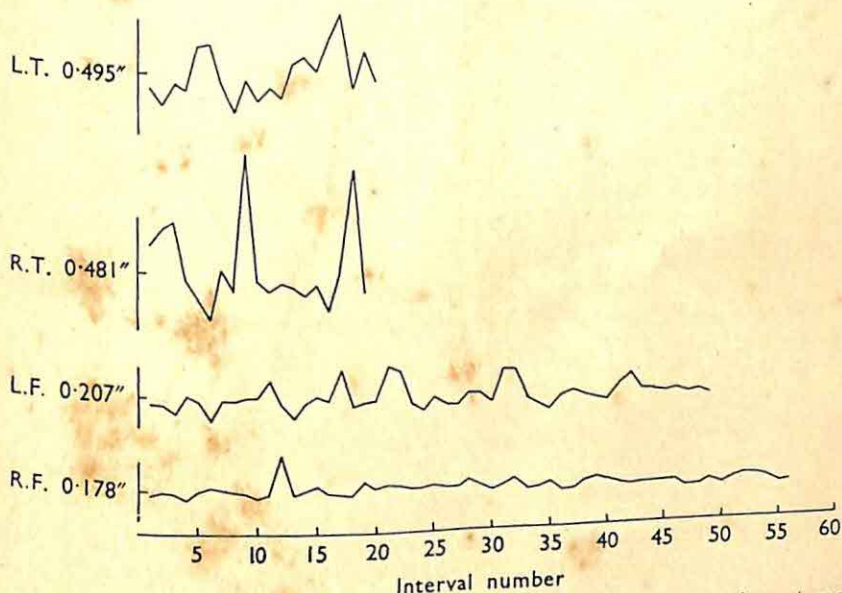
TABLE V

COMPARISON OF THE MEAN ERRORS (IN MMS.) OF THE RIGHT AND LEFT SIDES IN 10 SUBJECTS IN THE REPRODUCTION OF A PARTICULAR PRESSURE AND MAXIMUM RATE OF TAPPING (IN CYCLES PER SECOND) WITH THE INDEX FINGER

<i>Pressure</i> (fast application)		<i>Tapping</i>	
<i>Preferred</i> (right)	<i>Non-Preferred</i> (left)	<i>Preferred</i> (right)	<i>Non-Preferred</i> (left)
9.85	10.44	5.203	4.074
Not significant $t = 0.4297$		Significant $P < 0.01$	

Typical records are given in Figure 2, which shows the consistency of cycle length during runs with the fingers and toes of each side in one subject.

FIGURE 2



Variation of tapping interval on the two sides for one subject tapping at maximum speed with the index finger and big toe for ten seconds. The mean tapping interval (in seconds) is indicated in each instance. L = left side; R = right side; T = big toe; F = index finger.

There is a highly significant difference between sides in the aiming task, as can be seen in Table VI.

TABLE VI
COMPARISON OF THE MEAN ERRORS (IN CMS.) OF
THE RIGHT AND LEFT SIDES IN 10 SUBJECTS IN
AN AIMING TASK

<i>Preferred (right)</i>	<i>Non-Preferred (left)</i>
7.64	11.49
Significant $P < 0.01$	

DISCUSSION

The evidence from the second series of experiments confirms the hypothesis that differences between the two sides are more easily explicable in terms of previous usage or training than in terms of any inherited characteristic. Thus, there is no reliable difference in performance between the big toes of the two sides no matter what the type of task, presumably because one side has had no more opportunity for training than the other. Differences do show up in the finger-tapping and aiming tasks, probably because one side has received greater training either deliberately or incidentally in such activities. It may well be, of course, that with a much larger number of subjects, differences between the right and left sides in the task requiring accuracy of application of pressure would become evident, although this is hardly indicated in the present experiments.

It therefore seems true to say that handedness is at least mainly due to selective training and usage of one particular side. On the evidence from the present experiments, differences between the two sides only occur where the "timing" of the component parts of the movement are likely to be important, and where training could previously have taken place. The implication is, therefore, that this factor of timing is important in the learning of any movement where serial muscle activity is involved.

In a previous investigation where the maximum rate of turning a crank handle with the preferred and non-preferred hand was measured (Provins, 1953), a significant difference was found between the two sides in favour of the preferred hand. Furthermore, during the course of the experiment, considerable improvement in rate of turning was recorded by all subjects, who showed a greater mean improvement on the non-preferred side. At different phases of this investigation, the forces exerted at the control handle during the movement were recorded and are now analysed and discussed in the present report as the third series of experiments.

SERIES 3

INTRODUCTION

The question of "timing" has been discussed widely in the field of sport, and recently for example, in the official M.C.C. Cricket Coaching Book (1952) with special reference to fast bowling. It has also been stressed in golf (Murray, 1953) where particular attention has been paid to analysing the factors involved in the golf swing. In each instance emphasis is laid upon the correct timing of the various movements relative to one another. Electromyographic studies of the golf swing and tennis stroke (Slater-Hammel, 1948; 1949) confirm the preciseness of timing of the individual muscle contractions and their consistency in skilled players, but their wide variation between players.

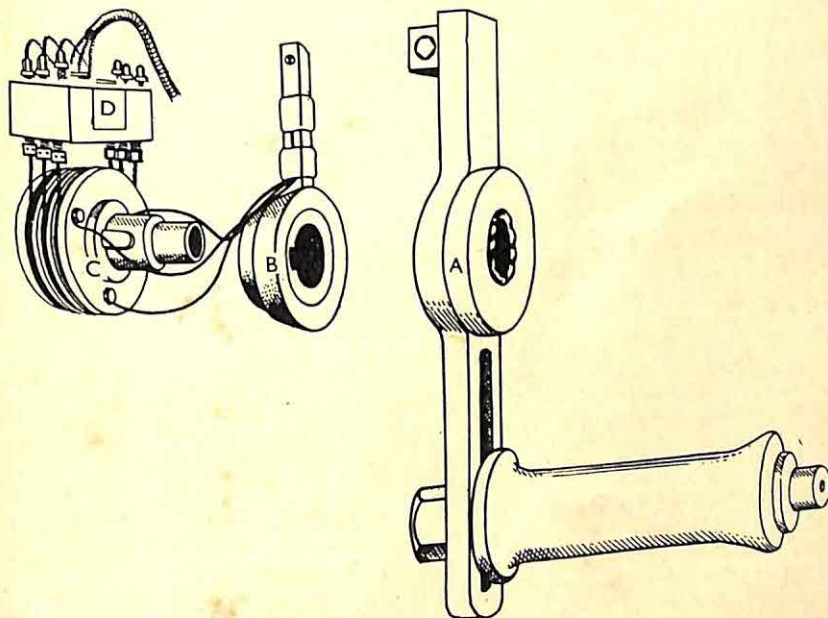
However, there appears to be no evidence available to show if this preciseness and consistency of timing improves with practice, and the present report was therefore undertaken with this aim in view.

In the finger-tapping and toe-tapping experiments described in series 1 and 2, strain gauge records were analysed in terms of consistency of tap interval and conclusions drawn concerning the timing of the alternate muscle contractions involved. In the present series, the same procedure was adopted using the records of force applied during a rotary movement of the hand in a handle-turning task which primarily involved the contraction of the flexors and extensors of the elbow and shoulder joints in a particular sequence. It was considered that measuring the time interval between corresponding muscle contractions in successive revolutions of the handle from force records taken at the beginning and end of the experimental series would provide a measure of any change in consistency of timing attributable to training. It should be pointed out that the experiment was primarily designed to study the effects of certain variables on the speed of turning a crank handle, and that the training which took place was incidental to the main investigation.

APPARATUS

The apparatus consisted essentially of a machine which mounted a crank and was capable of being adjusted to provide the required load or turning resistance. The working part of the machine consisted of an axle having a $4\frac{1}{4}$ -in. radius crank at one end and a polished metal drum of 3 in. radius keyed to the other. The drum rotated over a webbing belt, one end of which was secured to the frame of the machine and the other to a freely suspended 10 Kgm. weight, which was changed from one end of the belt to the other when the direction of rotation of the crank was reversed. The crank was mounted on two ball bearings and free to rotate on the shaft of the machine, but in fact turned it by driving the

FIGURE 3



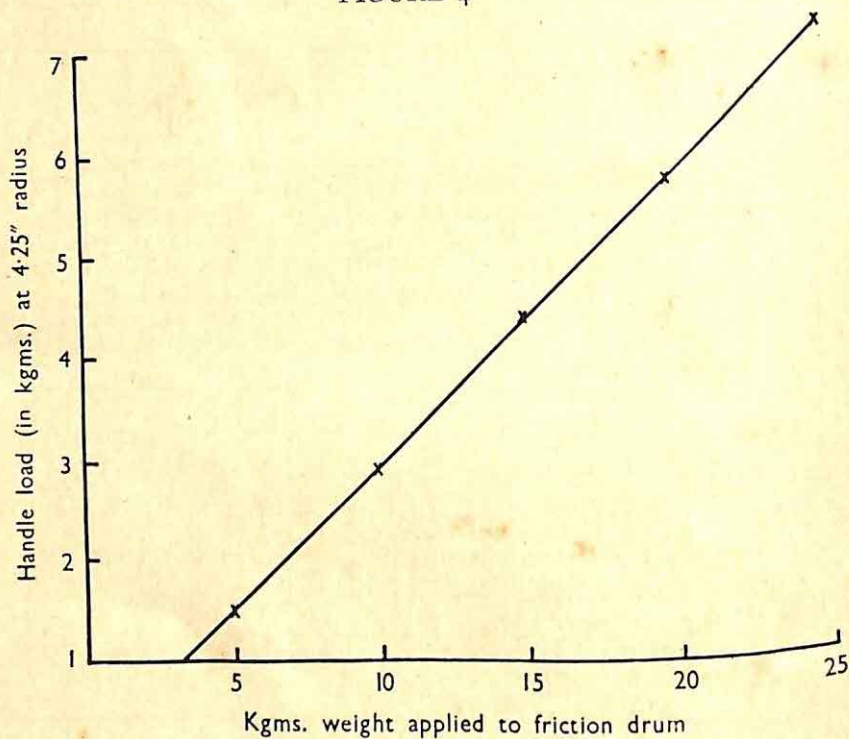
Strain-gauge crank

- (A) Crank fitted with ball bearing.
- (B) Keyed beryllium-copper unit with strain gauges.
- (C) Slip rings rotating with shaft.
- (D) Fixed ebonite block and wire brushes.

beryllium-copper unit which was keyed to the shaft as shown in Figure 3. The contact between the crank and the beryllium-copper bar on which were fitted two strain gauges, was by means of two steel balls which allowed movement of the two units with respect to one another in all directions except in the direction of rotation. The leads from the strain gauges were taken to three rhodium-plated slip rings which rotated with the shaft and on which three plaited silver-palladium wires made continuous contact. To the framework of the cranking machine was attached an ebonite block which mounted the wire brushes from which leads were taken to the rest of a Wheatstone bridge circuit and through an A.C. amplifier to a Brush pen recorder.

The apparatus was calibrated by connecting the machine through a flexible coupling to a variable speed D.C. motor. The crank was replaced by one-half of a flexible coupling which was driven by the motor against the load imposed by different weights hung on the webbing belt. The calibration graph of the apparatus is given in Figure 4. The torque at any braking resistance within the range indicated was found to be constant irrespective of the speed of rotation within the limits of the speeds used (25-250 r.p.m.).

FIGURE 4

*Calibration graph of cranking apparatus*

The frequency response of the recording system was limited by the frequency response of the amplifier and pen unit, which, however, was flat from 0.5-80 c/s.

METHOD

The apparatus was set up so that the axle of the machine was positioned transversely to the operator who was instructed to adopt and maintain as well as possible during the task, a comfortably erect sitting position. A typist's chair was used and a restraining belt fitted round the subject's waist to the bottom of the backrest prevented him grossly altering his position during work, but otherwise allowed him freedom of posture.

Six healthy male students were used as subjects and attended at the same time each day for five days a week for just under four weeks. In each session, each subject was required to turn the crank 54 revolutions at his maximum possible speed at each attempt and was given either eight or twelve attempts in all. The time taken was noted and the subject given knowledge of results and instructed to sit quietly and rest until his next turn after five minutes interval.

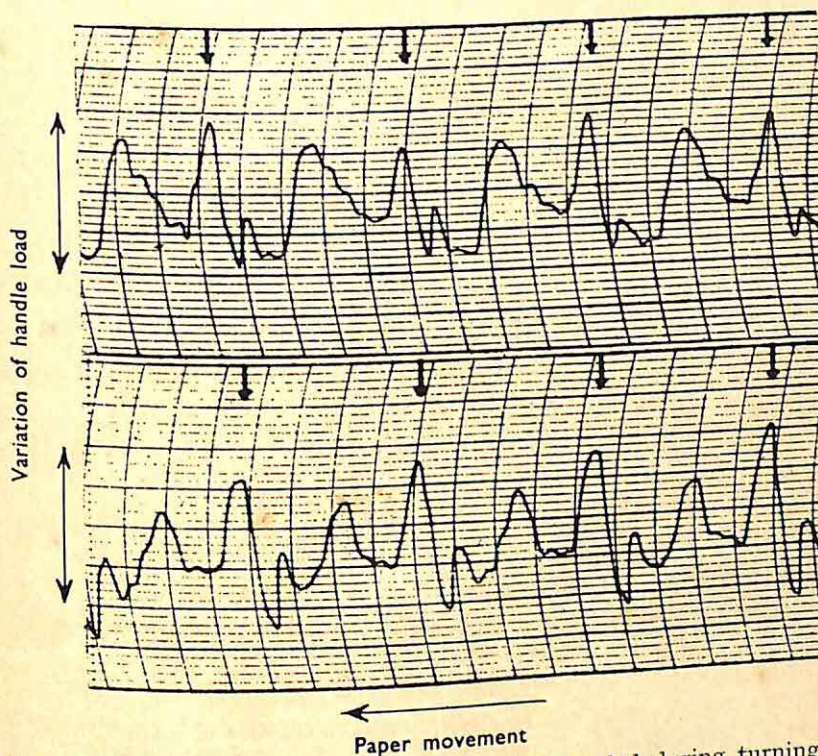
The experimental variables studied included seat height, reach, direction of rotation, hand used and training, as reported elsewhere (Provins, 1953).

At the end of the second session and the final session, the variation in force exerted at the handle of the crank during a typical run was recorded with the strain-gauge apparatus described. Between these two recordings each subject had received 176 practice attempts; 88 with each hand, 44 turning clockwise and 44 anticlockwise, in six different seat positions relative to the machine, representing the full range of working conditions in vertical height and forward reach.

RESULTS

Typical records at the end of training from two different subjects using the same hand turning in the same direction in approximately the same posture are given in Figure 5. The recording conditions were precisely the same for each record, which indicates the variation in force applied during each revolution, in consecutive revolutions.

FIGURE 5



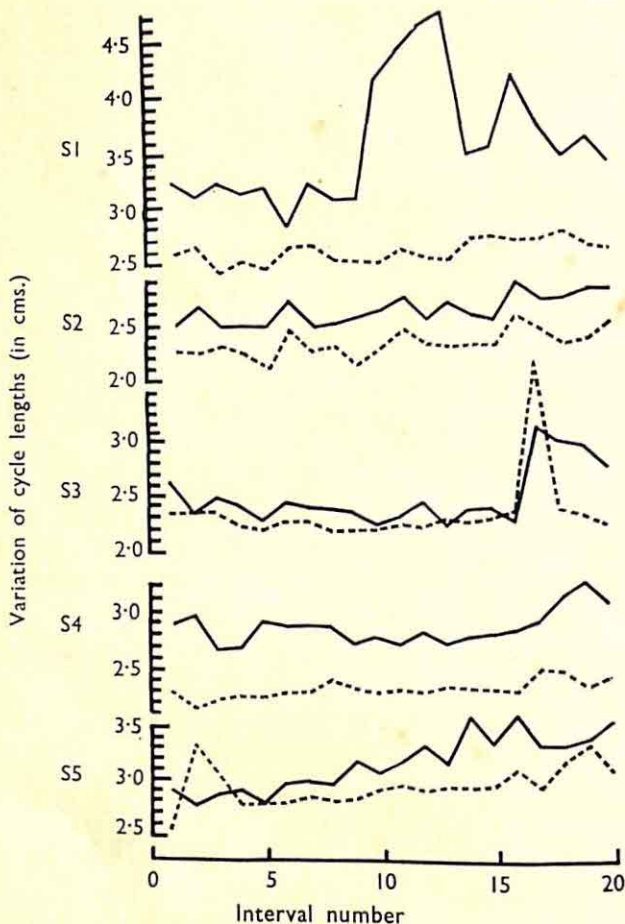
Variation in force applied to the crank handle recorded during turning with the non-preferred hand in two subjects (10 mms. corresponds to 4.5 lbs. variation in handle load; paper speed = 12 cms./sec.). Arrows indicate same position of crank handle in successive cycles.

It will be seen that while the general pattern of applied force is quite consistent for each subject, there are characteristic differences between the two records, which however, may have been due to slight differences in posture.

The records were analysed by measuring the distance between a characteristic well-defined peak in 20 successive revolutions and plotting the variations in cycle length as in Figure 6. Records of consistency are shown for five of the six subjects turning with the left hand in a clockwise direction at two different stages of training, just over three weeks apart. Records for the sixth subject are not available.

Greater consistency is evident after training in four of the five subjects, although this is more marked in some subjects than others. The distance between peaks gives a measure of the rate of turning which in every instance is, of course, higher after training. It is also interesting to note that the cycle length tends to increase during the run before training, whereas, after training, the inter-peak interval is much more constant.

FIGURE 6



Variation of interval (measured in cms.) between corresponding successive peaks in cranking records of five subjects before and after training (continuous line—before training, broken line—after training; paper speed 12 cms./sec.). The left-hand turning clockwise was recorded in each instance. Subject number 2 was left-handed, others right-handed. S = Subject.

DISCUSSION

While the difference in consistency between the cranking records for each subject before and after training is not always so marked as in the finger-tapping records between the preferred and non-preferred hand, it is perhaps surprising that the improvement shown in Figure 6 should be so clear as it is. The cranking experiment was not designed to provide a training situation and the frequent changing of seat position relative to the machine enforcing different body postures could not have been conducive to consistency in muscular activity. Even so, there seems to be a

definite improvement with training in the regularity with which a particular force was applied at the appropriate phase of the revolution. In other words, improvement in speed of cranking corresponded to an increase in consistency of the "timing" of the appropriate muscle contractions.

The importance of timing of the overall response relative to some stimulus, or in other words, co-ordinating receptor and effector functions has previously been demonstrated by Bartlett (1947) and others (Poulton, 1952; Singleton, 1953; Conrad, 1955). This presents a somewhat different aspect of skilled behaviour and concerns the appropriate moment at which the response as a whole should be made in relation to some external stimulus and may be considered as primarily a perceptual problem. There are thus two stages involved, one concerning judgment of the appropriate moment for the movement to be made in relation to external conditions, and the other (which has been discussed in this paper) concerning the serial organization of each muscle contraction comprising the movement. The timing of the response relative to some stimulus is therefore dependant on the organization of the response, in the same way that the organization of the response or timing of the individual muscle contractions must be dependant on knowledge of the appropriate muscles to contract in order to effect the required movement. As Fleishman (1953) points out . . . "the distinction between perceptual and motor skills is a somewhat arbitrary one, each class of skills being represented in varying degrees in the performance of different tasks." The bomb-aimer, quoted by Bartlett (1947) has mainly a perceptual task, whereas the dart-player or golf-player has primarily to master a highly complex movement to achieve much success. Most skilled tasks lie between these two extremes and consequently demand a certain degree of organization at both levels of behaviour. As Welford (1951) points out. . . . "It can thus fairly be said that there is in use, at any time in the performance of a skilled task, not one but a whole hierarchy of organizations. Some of these are concerned with the immediate details of what is being done at the moment; others are concerned with more general aspects of the task, integrating the details into larger units—in a sense *using* them."

CONCLUSIONS

The evidence suggests that: Firstly, handedness is determined largely by training and environmental factors rather than by any inherited characteristics. Certainly, no differences were found in the performance of corresponding parts of the body on the two sides where no previous training could have taken place, but only where previous training was possible. The amount of difference where it occurred could easily be explained by the extent of differential training on the two sides, as, for example, the relatively small difference shown in the aiming task for the ten females as compared with the highly significant difference between the two sides in the same task where nine of the ten subjects were males. Secondly, if skill is defined as "ability gained by practice" (Thorndike, 1948), then one of the essential characteristics of a skilled movement is the "timing" of the various muscles used. This "timing" or "phasing" of muscular contractions is displayed in any movement where several muscles act serially or alternately.

ACKNOWLEDGEMENTS

I am grateful to Mr. R. M. Jones for his advice on statistical matters, and to Mr. G. Toothill, of the Military College of Science, and to Mr. E. Reeves for the design and construction of much of the electronic apparatus.
I wish to thank Miss E. M. Macdonald and the staff and students of the Dorset House School of Occupational Thereapy for their help and co-operation in these experiments.

REFERENCES

- BARTLETT, F. C. (1947). The measurement of human skill. *Brit. Med. J.*, **1**, 835-838; 877-880.
- BLAU, A. (1946). *The Master Hand*. New York.
- BRAIN, R. (1945). Speech and handedness. *Lancet*, 837-841.
- BURGE, I. C. (1952). Some aspects of handedness in primary school children. *Brit. J. educ. Psychol.*, **22**, 45-51.
- COLE, J. (1955). Paw preference in cats related to hand preference in animals and man. *J. comp. physiol. Psychol.*, **48**, 137-140.
- CONRAD, R. (1955). Adaptation to time in a sensorimotor skill. *J. exp. Psychol.*, **49**, 115-121.
- CROMWELL, H., and RIFE, D. C. (1942). Dermatoglyphics in relation to functional handedness. *Hum. Biol.*, **14**, 516-526.
- DUROST, W. H. (1934). The development of a battery of objective group tests of manual laterality with the results of their application to 1,300 children. *Genet. Psychol. Monog.*, **16**, 225-335.
- FLEISHMAN, E. A. (1953). Testing for psychomotor abilities by means of apparatus tests. *Psychol. Bull.*, **50**, 241-262.
- GIBBS, C. B. (1954). The continuous regulation of skilled response by kinaesthetic feedback. *Brit. J. Psychol.*, **45**, 24-39.
- GLEES, P., and COLE, J. (1952). Ipsilateral representation in the cerebral cortex. *Lancet*, **1**, 1191-2.
- GOODGLASS, H., and QUADFASSEL, F. A. (1954). Language laterality in left-handed aphasics. *Brain*, **77**, 521-548.
- GRANT, D. A., and KAESTNER, N. F. (1955). Constant velocity tracking as a function of subjects' handedness and the rate and direction of the target course. *J. exp. Psychol.*, **49**, 203-208.
- HILDRETH, S. H. (1949a). The development and training of hand dominance. 1. Characteristics of handedness. *J. genet. Psychol.*, **75**, 197-220.
- HILDRETH, S. H. (1949b). The development and training of hand dominance. 2. Developmental tendencies in handedness. *J. genet. Psychol.*, **75**, 221-254.
- HILDRETH, S. H. (1949c). The development and training of hand dominance. 3. Origins of handedness and lateral dominance. *J. genet. Psychol.*, **75**, 255-274.
- HILDRETH, S. H. (1950a). The development and training of hand dominance. 4. Developmental problems associated with handedness. *J. genet. Psychol.*, **76**, 39-100.
- HILDRETH, S. H. (1950b). The development and training of hand dominance. 5. Training of handedness. *J. genet. Psychol.*, **76**, 101-144.
- HOLT, S. B. (1954). Genetics of dermal ridges: bilateral asymmetry in finger-ridge counts. *Ann. Eugen.*, **18**, 211-231.
- JACK, W. R. (1895). On the analysis of voluntary muscular movements by certain new instruments. *Proc. roy. Soc.*, **57**, 477-481.
- M.C.C. (1952). *The M.C.C. Cricket Coaching Book*. London.
- MURRAY, H. A. (1953). *The Golf Secret*. Kingswood.
- NEWMAN, H. H. (1934). Dermatoglyphics and the problem of handedness. *Amer. J. Anat.*, **55**, 277-322.
- POULTON, E. C. (1952). Perceptual anticipation in tracking with two-pointer and one-pointer displays. *Brit. J. Psychol.*, **43**, 222-9.
- PROVINS, K. A. (1953). A study of some factors affecting speed of cranking. Report to the Medical Research Council. R.N.P. 53/755.
- PROVINS, K. A. (1955). The rôle of receptors in muscle and tendon in controlling the application of finger pressure in man. *J. Physiol.*, **128**, 55-6P.
- RIFE, D. C. (1943). Handedness and dermatoglyphics in twins. *Hum. Biol.*, **15**, 45-54.
- RIFE, D. C. (1950). An application of gene frequency analysis to the interpretation of data from twins. *Hum. Biol.*, **22**, 136-145.
- SINGLETON, W. T. (1953). Deterioration of performance on a short-term perceptual-motor task. In Floyd, W. F., and Welford, A. T. *Symposium on Fatigue*. London.
- SLATER-HAMMEL, A. T. (1948). Action current study of contraction movement relationships in the golf stroke. *Res. Quart. Amer. Assoc. Hlth. Phys. Educ. & Rec.*, **19**, 164-177.
- SLATER-HAMMEL, A. T. (1949). Action current study of contraction movement relationships in the tennis stroke. *Res. Quart. Amer. Assoc. Hlth. Phys. Educ. & Rec.*, **20**, 424-431.

- THORNDIKE, E. L. (1948). *Thorndike English Dictionary*. London.
WELFORD, A. T. (1951). *Skill and Age*. London.
ZANGWILL, O. L. (1954). Agraphia due to a left parietal glioma in a left-handed man.
Brain, 77, 510-520.

BOOK REVIEW

The Subnormal Mind. By Sir Cyril Burt. Third Edition. London: Geoffrey Cumberlege. Oxford University Press. 1955. Pp. 391 + xix.

This is a third edition of the book which first appeared in 1935. It has a new preface and footnotes, but there appear to be no important changes in the text. The book describes and explains all the mental abnormalities of children, intellectual and temperamental, with the exception of childhood psychosis. In the last twenty years it has become widely recognized that psychotic conditions can exist in childhood and it is perhaps a pity that Professor Burt has not brought his book up to date in this respect. Apart from this, intellectual and personality abnormalities in children are discussed comprehensively and with case histories and as description the account is probably the best available.

In psychology it is refreshing to read a book which is written judiciously and with common sense. Professor Burt does not put forward any views which might seem to the layman implausible or dangerous. He advocates firm discipline for the delinquent, reassurance for the anxious, special schools for the backward. The attitude is broad-minded and eclectic. While viewing psychoanalytic theory with caution, the author accepts the Freudian mechanisms accounting for phobias and conversion hysteria. Due allowance is made for both constitutional and environmental factors in the discussions of the causation of mental conditions.

Theoretically, Professor Burt rests largely upon the theory of instincts and sentiments put forward by McDougall. Some readers, especially if they are experimental psychologists, may regard the foundations of this theory as insubstantial. Probably the chief problem concerns the differentiation of the postulated instincts and the implication that empirical relationships can be established between stimulus and response. Accepting the dualism in which this theory is cast, in the case of both the observable and the emotional response a distinction should be made between the social instincts and those induced by some internal condition (e.g. food seeking, sex), together with escape. For the second group, relationships can be established between stimulus and response, so long as the response is defined as an activity which terminates the stimulus situation, e.g. relationships can be established between hours of food deprivation and eating activity. It is doubtful whether any comparable relationships can be found for the social instincts; one of the most impressive attempts to establish such a relationship, the frustration-aggression hypothesis, has been recognized as untenable.

A similar distinction should be made in the question of the emotional response. It is likely that in the case of internally induced emotions relationships can be established between stimulus conditions and conscious feelings, e.g. food deprivation and feeling hungry. Everyday experience suggests that no such inductions are possible for externally-induced conditions. Thus given a policeman's rebuke, some people will feel fear, some anger, and others pride, curiosity, gregariousness or mirth. For this reason it seems plausible to many writers to regard these reactions as *learned*. At any rate, in the present state of knowledge there is little evidence to support Professor Burt's view that delinquency should be ascribed, even in part, to a strong instinct of pugnacity, or even that such an instinct exists.

These theoretical issues, however, do not greatly affect the descriptive content of Professor Burt's book and this is its chief purpose. As a straightforward account of the psychological disorders of children it will remain a standard work.

R. L.

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY

Vol. VIII

1956

Part 3

THE PHYSICAL PROPERTIES OF BAR-PRESSING BEHAVIOUR AND THE PROBLEM OF REACTIVE INHIBITION

BY

J. R. TROTTER

From the Institute of Experimental Psychology, University of Oxford

Many of the terms used in the literature on reactive inhibition are vague, subjective or incorrect. The weighted bar is not a suitable device in the study of this problem, since it records irrelevant aspects of behaviour. An experiment using apparatus which gives a continuous record of the force which a rat applies to a knob produces the following results. After much practice, pressing becomes sharp and brief and the amount of activity per reward is reduced. Under conditions of no (intentional) secondary reward the amount of activity during extinction is at first positively correlated with the average activity per reward during training, but the correlation diminishes as extinction proceeds. With auditory secondary reward there is no correlation.

INTRODUCTION

Numerous experiments have been performed to test the reactive-inhibition hypothesis that responses involving more effort are more easily extinguished. The experiments considered here are those in which a weighted bar has been used. A typical treatment of the problem is that of Mowrer and Jones (1943). They trained rats to press a bar for food. Weights could be fixed to the far end of the bar to vary the effortfulness of moving it. The weights on successive trials were 5, 30, 5, 55, 5, 80, 5, 80, 5 and 80 gm. The final rates of pressing the bar weighted with 5 gm. and with 80 gm. were 5 per min. and 4.1 per min. and equality of training on the two conditions was inferred. For extinction trials the rats were randomly assigned to three groups and presented with the bar weighted with 5 gm., 42.5 gm. and 80 gm. The mean numbers of responses during extinction were 350.3, 248.4 and 110.7, this result being taken to indicate a negative correlation between effortfulness of task and resistance to extinction, and to support the hypothesis.

This experiment is mis-quoted, to the effect that training was equal on all three weights, by Hull (1943), Solomon (1948a and 1948b) and Osgood (1953), but these authors accept the conclusion. Applezweig (1951), Montgomery (1951) and Thompson (1944) have investigated related problems using similar apparatus. Maatsch *et al.* (1954) criticized these experiments and concluded that work "... bears no relation to resistance to extinction". This statement is offered as a judgment of fact, not as a logical principle. Stanley and Aamodt (1954) agree with this but consider that the Mowrer and Jones study "... may be regarded as demonstrating that requiring forceful responses during conditioning yields forceful responses during extinction."

There are three types of variable referred to in this literature:—(1) variables like effort, difficulty and fatigue, rich in associations which give the hypothesis a *prima facie* plausibility, but vague and subjective, (2) various (putative) physical correlates of effort, such as "... energy consumption or work (W)" (Hull, 1943) or gm. mass (Stanley and Aamodt, 1954), which are mentioned but not measured, and (3) the experimental independent variable weight, which ostensibly determines the effortfulness of pressing the bar, and the experimental dependent variable, number of responses. Much confusion can result if discourse is allowed to slip from one type of term to another. The terminology preferred by the writer is given below.

Energy expenditure and mechanical work

Power is expended whenever a muscle is activated. If instantaneous power varies, the total energy expended is the time integral of power. Under isometric conditions instantaneous force varies with power, and impulse, the time integral of force, is correlated with energy, but no external work is done. When the muscle is activated without constraint, energy is expended, impulse is nil, and no external work is done. In both cases all the energy generated is dissipated as heat. Though no work is done, prolonged stimulation leads to a decrement in activity. In the general case, when the muscle contracts to some extent against an opposing force, part of the energy generated is devoted to mechanical work outside the muscle.

Energy expenditure and impulse are functions of force and *time*; work is a function of force and *displacement*. There is no general relation between energy and work; the relation depends on the skeletal and mechanical systems to which the muscle is attached. In the intact animal, impulse is an approximate measure of energy expenditure, provided that movement is slow and all presses are performed in a similar manner. Under the same conditions force is an approximate measure of power.

There are two important points:—(1) work *per se* is irrelevant to any study of fatigue or related phenomena, (2) if energy expenditure is the relevant factor, work is not an accurate measure of this.

The weighted bar

Consider a pivoted bar of negligible mass, loaded at the far end with a weight W gm. at a mean distance L cm. from the pivot. The bar can rotate a small angle about the horizontal, between two stops. The weight applies a moment of force $W \times L$ gm. cm. to the bar. At a point l cm. from the pivot at the near end the force required to move the bar and produce a reward (F_R) is determined by the relation $F_R \times l \geq W \times L$. If the region of the bar where force may be applied is fixed, and small compared with F_L , $F_R \geq W \times L/l$. The last condition is essential if weight is to be a valid experimental variable, i.e. to set the rat a specific situation. Mowrer and Jones (1943) state that their bar projected $4\frac{1}{2}$ in. into the cage. Judging by a photograph of their apparatus, all parts of the $4\frac{1}{2}$ in. appear more or less accessible to the rat, and since the pivot appears to be about 1 in. beyond the cage wall, a variation of leverage of the order of 5 to 1 is possible. If the weights determined the force required at the near extremity of the bar, then the weights used, 5 gm., 42.5 gm. and 80 gm. set ranges of F_R of the order of (5–25) gm., (42.5–212.5) gm. and (80–400) gm.

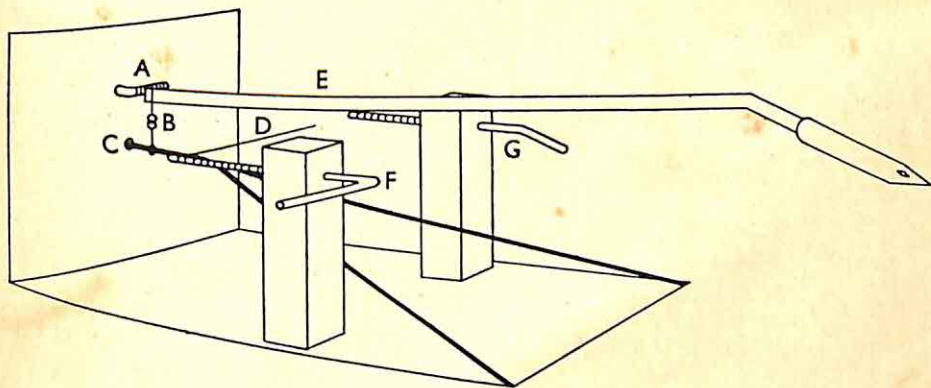
(a) When the applied force is less than F_R gm., no matter how long the force is sustained, the bar does not move. Impulse and energy are expended, but the recording apparatus, which depends on movement, makes no record. No work is done on the bar.

(b) If the applied force only slightly exceeds F_R gm. the bar moves slowly to the stop, a distance of s cm. A reward is delivered and one "response" is recorded, which it is conventional to regard as one unit of effort. It is clear from the discussion above that such a "response" cannot be regarded as a consistent unit of impulse or of energy. An amount of work $W \times s$ gm. cm. is done on the bar, giving it potential energy.

(c) If a force greatly exceeding F_R gm. is applied, the bar moves rapidly and hits the stop with velocity v . Work done on the bar equals $W \times s$ gm. cm. potential energy plus $W \times v^2/2g$ gm. cm. kinetic energy. The presence of the kinetic energy term, which is unpredictably variable, prevents the adoption of a "response" as a consistent unit of work. The variable weight is presumably chosen because of a supposed linear relation with work. There is no such relation, and even if there were, work itself is irrelevant.

The important variable of which this apparatus takes no account is the duration of a press. Unless this is considered, *no* quantitative account of pressing behaviour is possible. The calculations in the experiments quoted are based on the assumptions that each recorded "response" to a given weighted bar involves one unit of "effort," and that the size of the unit is determined linearly by the weight on the bar. For any reasonable physical interpretation of "effort" these assumptions are groundless.

FIGURE 1



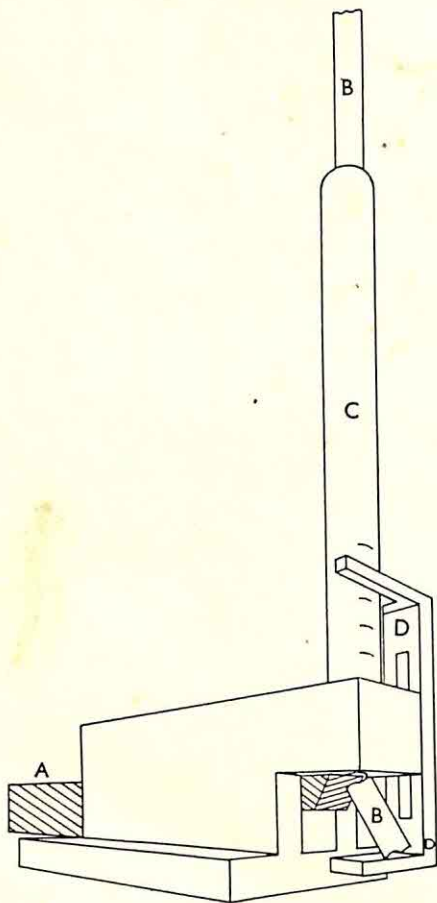
The force recorder.

The writing arm is hinged on a spring pivot (A), and is connected by a jointed coupling (B) to the V-shaped beam. The rat presses a knob on the end of the beam, beyond the hole (C). Wire contacts (D and E), raised or lowered by the handles (F and G), when touched by the writing arm, operate and reset the spaghetti gun at any desired force. The knob moves downwards about 5.5 microns per gm. applied force, moving the pen about 1 cm. per 50 gm.

This account casts doubt on the usefulness of some classifications of responses. Stanley and Aamodt (1954) refer to "force requirement per response," determined by the weight on the bar, and they used two classes of response, *incomplete* bar presses, i.e. presses between 1 mm. and 1 cm., and *complete* bar presses of 1 cm. This distinction is irrelevant to impulse or energy. Applezweig (1951) employed three categories of responses, partial pressings and pressings (equivalent to the above two classes) and holding responses "... those in which the lever is depressed and not released." Skinner (1938) presents results in terms of "grams per response." Presses described under (a) above are nowhere considered.

It is evident from mechanical principles that any expression equivalent to "the force of a press" is ambiguous or syntactically incorrect. It may mean, either a variable property of a press (f), or mean force during the active time (f_a), or peak force (p) or any other function, e.g. R.M.S. force, which has the requisite dimensions. (See Fig. 3.) Force does not characterize a press-as-a-whole; impulse does this.

FIGURE 2



The Deutsch spaghetti gun.

In the forward position the blade (A) supports the spaghetti (B) in the feed tube (C). When retracted against a spring by an electro-magnet, the blade breaks the magnet circuit, and on returning cuts the spaghetti which has fallen a measured distance to the adjustable platform (D).

THE EXPERIMENT

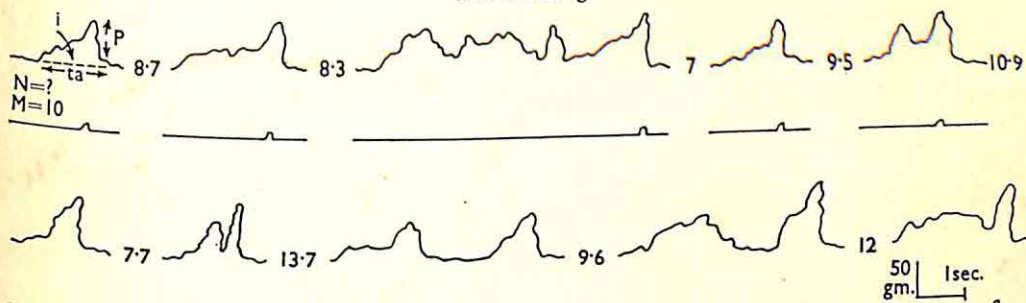
The questions treated in the present experiment are here stated in informal terms. It is found that in various situations different rats devote different amounts of activity, on the average, to obtaining a single reward. How does the amount and kind of activity vary during practice? How is the amount of unrewarded activity during extinction related to the average amount expended per reward?

Apparatus

The rat is trained to press the knob of the force recorder in order to obtain a reward of 5 mm. of spaghetti. The knob is a $\frac{1}{2}$ -in. sphere placed about 2 in. from the floor of the cage. Only the knob is available to the rat, and it is fixed to the end of a 4-in.

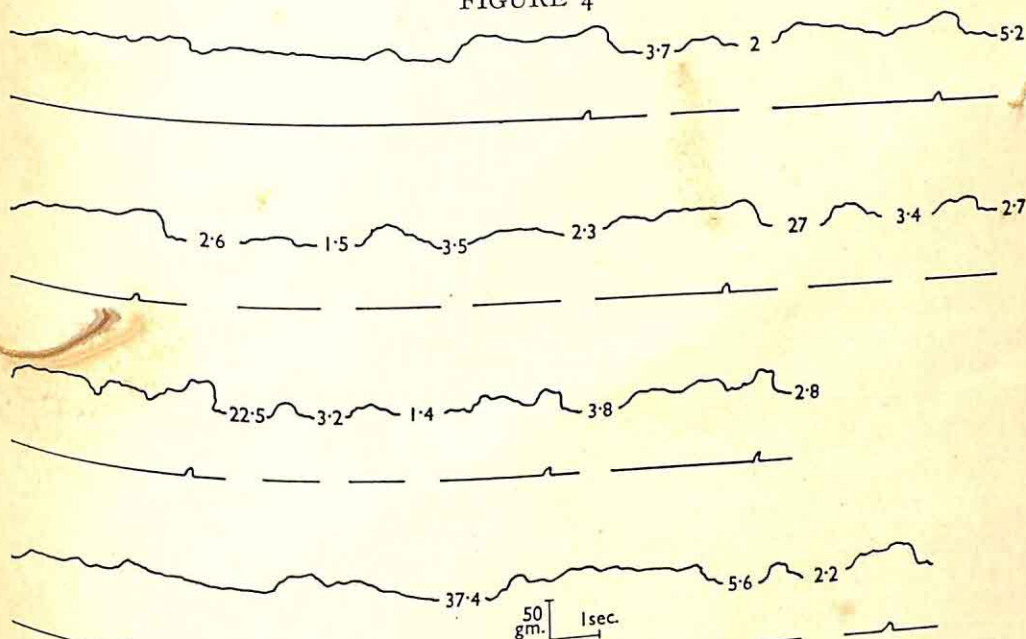
pivoted bar, so the possible variation of leverage is negligible. The force recorder, illustrated in Figure 1, makes a magnified inked record of the movement of the knob, on a paper moving at about 1 cm. per sec. The bar is not weighted but sprung, so movement is smooth and proportional to the force applied. When a force of F_R gm. (a value chosen by the experimenter) is applied, the force recorder operates the Deutsch spaghetti gun which delivers a piece of spaghetti, and a pen records the event. The spaghetti gun is illustrated in Figure 2.

FIGURE 3



Training record of a well trained rat.

FIGURE 4



Training record of a poorly trained rat.

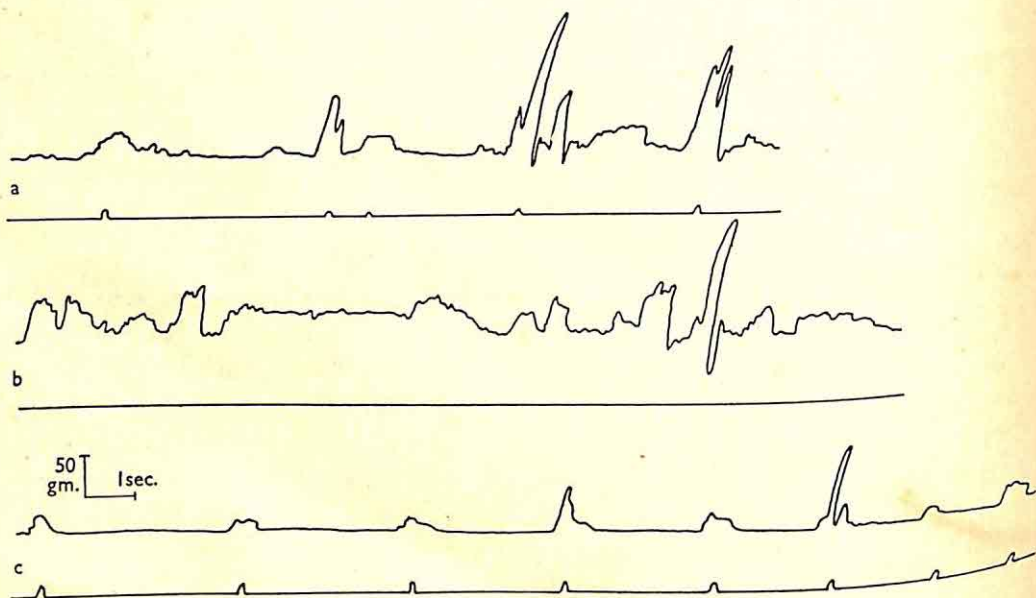
Some records are shown in Figures 3, 4 and 5. The record is read from left to right, with force indicated by the height of the graph, and time indicated horizontally. Where the original record indicated that the knob was undisturbed for some time, the portion indicating zero force was removed, and the inserted number indicates its duration in seconds. The kinks in the lower line indicate operation of the spaghetti gun. Some records, like Figure 3, consist of well-defined portions separated by intervals of zero force. It is convenient to regard these isolated portions as due to discrete presses on the knob. In Figure 4 well-defined presses are not so evident. Figures 5a and 5b do not invite division in any systematic manner.

Method

The subjects were 12 male hooded rats aged about six months. Parts of seven days were spent in training each rat with F_R at less than 5 gm. During the next five days each rat was allowed to produce 25 rewards daily. The rats were fed together for one hour after the completion of all daily trials, so spent upwards of 20 hours without food before the next trial. At the end of this preliminary training all rats had a well-established knob-pressing habit. Though no systematic records were taken up to this time, it was observed that a force exceeding 10 gm. was seldom applied.

For the experiment, four rats were randomly assigned to each of the 10 gm., 30 gm. and 50 gm. groups. For all groups F_R was initially 10 gm. The 10 gm. group continued at this value. For the 30 gm. group 5 gm. increments were made to F_R every other day up to a maximum of 30 gm. For the 50 gm. group 5 gm. increments were made to F_R every day up to a maximum of 50 gm. 25 rewards were allowed daily. For the next four days daily trials of 15 rewards were made at the maximum value of F_R . Thus at the end of training the three groups had well-established habits with F_R at 10 gm., 30 gm. and 50 gm.

FIGURE 5



Extinction records.

For the extinction trials each group was divided into a *with* class and a *without* class, two rats in each. Rats in the *with* class were allowed to hear the sound of the empty spaghetti gun when they pressed with the maximum value of F_R . The sound is loud and distinctive. The number of successful presses (presses reaching F_R gm.) was recorded. For the *without* class the spaghetti gun and its recording pen were disconnected. The successive daily extinction trials were ten minutes, five minutes and three minutes.

Daily records were made during the first 15 rewards of the 50 gm. group from the 10 gm. trial to the last 50 gm. trial. Records of the last two training trials were made for the 10 gm. and 30 gm. groups. For all groups records were made throughout the extinction trials.

Data on the record and their utilization

The source of data is the record, which indicates a variation of force (f) through time (t).

When the record has been divided into presses (in a case where this procedure is feasible) the following variables can be measured, as illustrated in Figure 3. Number of presses (N) and number of rewards (M); active time (t_a), peak force (p) and impulse ($i = \int f \cdot dt$) for each press.

The gross characteristics of pressing behaviour during any period are measured by total active time ($T_a = \Sigma t_a$), total impulse ($I = \Sigma i$) and mean force during the active time ($F_a = I/T_a$).

The characteristics of an average press in any performance are average active time per press ($\bar{t}_a = \Sigma t_a/N$), average peak force per press ($\bar{p} = \Sigma p/N$) and average impulse per press ($\bar{i} = \Sigma i/N$). A measure of the form of an average press, independent of its absolute magnitude, is the form-ratio ($\bar{p}/\bar{t}_a$). A large form ratio indicates peaky pressing, and a small one indicates a gradual increase and decrease of force throughout the press.

Functions which measure the average amount of activity devoted to producing a reward in any performance are average active time per reward ($t'_a = \Sigma t_a/M$) and average impulse per reward ($i' = \Sigma i/M$).

Changes in pressing during training

Functions calculated for the 50 gm. group on each training trial are shown in Table I.

TABLE I
RECORDS OF THE 50 GM. GROUP ON SUCCESSIVE TRAINING TRIALS

Force required for reward	Trial		Average per press				Average per reward	
	Number of presses	Mean force during the active time	Active time	Impulse	Peak force	Form ratio	Active time	Impulse
F_R gm.	N	F_a gm.	$\bar{t}_a$ sec.	$\bar{i}$ gm. sec.	$\bar{p}$ gm.	$\bar{p}/\bar{t}_a$	t'_a sec.	i' gm. sec.
10	27.8	6.1	0.68	4.13	11.7	17.2	1.25	7.6
15	47.3	6.0	1.27	7.66	12.1	10.4	4.0	24.1
20	25.0	7.3	1.52	11.2	17.5		2.6	18.6
25	25.3	10.6	2.24	23.9	22.6		3.78	40.3
30	32.0	8.8	2.17	19.2	23.4		4.57	40.8
35	37.3	11.6	2.17	25.2	27.3		5.37	62.5
40	19.0	15.2	2.14	32.4	40.1	28.1	2.7	41.1
(1) 45	20.5	14.5	2.91	42.6	48.6		4.0	58.2
(2) 50	25.3	21.5	2.06	44.3	49.3		3.48	74.8
(3) 50	20.5	20.7	1.84	38.3	51.3		2.54	52.5
(4) 50	16.5	19.1	1.98	36.8	59.0		2.17	41.6
(5) 50	21.5	23.1	1.41	32.5	52.2	28.1	2.02	46.6
50	17.3	21.0	2.57	54.1	55.4		2.95	62.0

The requirement in the 10 gm. trial did not greatly exceed the rats' accustomed peak force, and their performance was moderately skilled, compared to others in the experiment. If the next four trials (in which F_R was repeatedly raised, making accustomed behaviour inadequate) are compared with the last four trials (in which practice was obtained under one condition) the following trends are seen:—In the later trials the number of presses N approximates more closely to number of rewards M (15). The average active time per press is not greatly changed (1.87 : 1.74), despite the difference in F_R , the average peak force is increased to meet the greater F_R , and peakiness is increased. The average active time per reward is reduced.

There is more evidence on these changes. The 50 gm. group performed with F_R at 30 gm. on the fifth training trial, before performance was at its best. The 30 gm. group performed with F_R at 30 gm. on the last five training trials, when they had had much practice. The functions for the two groups are:—

Group	F_R	Trial		Press				Reward	
		N	F_a	$\bar{t}_a$	$\bar{i}$	$\bar{p}$	$\bar{p}/\bar{t}_a$	t'_a	i'
50 gm.	30	32	8.8	2.17	19.2	23.4	10.8	4.57	40.8
30 gm.	30	20.5	15.2	1.43	21.7	33.3	23.3	1.95	29.6

The corresponding figures for the 50 gm. group on the 10 gm. trial and the 10 gm. group on the last two training trials are:—

Group	F_R	Trial		Press				Reward	
		N	F_a	$\bar{t}_a$	$\bar{i}$	$\bar{p}$	$\bar{p}/\bar{t}_a$	t'_a	i'
50 gm.	10	27.8	6.1	0.68	4.13	11.7	17.2	1.25	7.6
10 gm.	10	19.3	6.3	0.78	4.42	13.9	17.6	1.0	4.92

The general trends are:—fewer unsuccessful presses, increased average force during the active time, briefer and more peaky pressing, and reduction in average active time per reward and average impulse per reward.

Now if a rat habitually presses with a small peak force and only seldom attains F_R gm. it will necessarily expend much time and impulse per reward, since so much of each is wasted. But the rat which consistently presses with a large peak force, which it does only after considerable training, quickly produces rewards and is able to cease pressing, thereby saving time and impulse. Thus any tendency towards "economy of effort" which expresses itself in a reduced rate of activity, i.e. reduced force and power, necessarily fails in its presumed purpose of reducing total expenditure. The reductions in t'_a and i' , far from substantiating "economy of effort" as considered above, indicate an opposed tendency, that the well-trained rat readily exerts itself.

Behaviour during extinction

A measure of performance during training is the average impulse expended per reward (i'), and a measure of extinction performance is the total impulse expended (I_E). These two variables are plotted in Fig. 6, for the *with* and the *without* classes separately. For the *without* class, on each separate period and for the total extinction time, the correlation coefficients and their probabilities are:—

Period	r	p
1	0.895	0.013
2	0.826	0.043
3	0.61	0.219
Total	0.87	0.022

Initial extinction performance is definitely correlated with training performance, but the correlation diminishes as extinction proceeds.

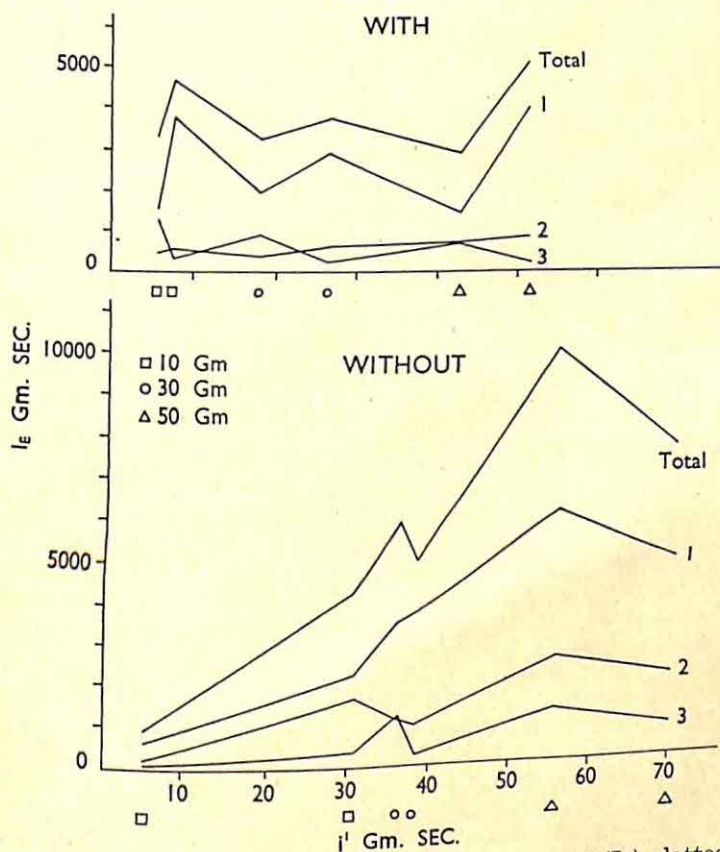
For the *with* class, which was given auditory "secondary reward," there is no correlation whatever. The writer reserves comment on this, since the problem of "secondary reward" is now being investigated more thoroughly.

Figure 5 shows the variability of pressing behaviour during extinction. The records are from three rats in the 10 gm. group. Peak forces exceeding 100 gm. are applied for the first time in the experiment. These rats had been required to exert only 10 gm. for reward, and had occasionally but not consistently exerted more.

But during extinction, rats of all groups, no matter what forces they had been accustomed to during training, applied forces which are probably near the limit of the animals' capacity. This negates the conclusion of Stanley and Aamodt (1954) regarding the correlation between training forces and extinction forces.

Judging by Figures 3 and 5, the rat, when well-trained, remains active until a reward is obtained and then stops pressing in order to eat. Behaviour is divided into portions by eating spells, and average eating time (about 10 sec. in Fig. 4) is the main factor determining rate of pressing. A press which does not immediately produce a reward is either prolonged or followed quickly by a further press. When the food apparatus is disconnected, no eating spells occur to interrupt pressing, so pressing occurs in longer spells (Figs. 5a and 5b) or is interrupted by short intervals (about 3 sec. in Fig. 5c).

FIGURE 6



Impulse expended during extinction periods 1, 2, 3 and total (I_E) plotted against average impulse expended per reward during training (i').

Output and result in the weighted bar situation

This experiment illustrates an important distinction which seems obvious when stated explicitly, but which is commonly obscured by the ambiguous use of the term "response."

A rat is put into a weighted-bar situation and its behaviour is studied. Situation variables such as weight on the bar and delay of reward are under the experimenter's control. That part of the rat's behaviour in which it exerts force on the bar may be called the output. It is measured by impulse. The result, which consists of events

produced in the situation by the output, includes movement of the bar, delivery of food, noise of mechanism, etc. The result is determined jointly by output and situation. Given a certain temporal variation of force as the output, the result in terms of number of bar movements, etc., will depend on the weight on the bar. If the weight is great, a small number of movements is no indication of a small amount of output.

Now response-produced (reactive) inhibition is evidently intended to mean output-produced inhibition. But the measure commonly used is number of "responses" (of rewards or bar-movements). Result, not output is measured. This shows that the simple weighted bar situation is not suitable for studying reactive inhibition.

The great variability of output in a constant situation is shown in Figure 6. The horizontal dispersion of the values for the 10 gm. group (squares), the 30 gm. group (circles) and the 50 gm. group (triangles) indicates the wide range of average impulse per reward (i') for a constant F_R . The use of F_R as an independent variable to be compared with output during extinction would implicitly equate all training outputs. The experimental situation here is different from the weighted-bar situation in that no abrupt movement of the bar accompanies delivery of a reward. It might be suggested that a greater uniformity of pressing occurs in the weighted bar situation, but work in hand negates this.

By the use of i' and I , output is compared with output. The conventional treatment consists of comparing a situation variable, weight on the bar, with a result variable, number of bar movements. The data of the present experiment may be treated in the conventional manner, since F_R is equivalent to the weight on the bar. The results for the *with* class are:—

F_R	10 gm.	30 gm.	50 gm.
Number	217	115	86

These figures are in numerical agreement with those of Mowrer and Jones (1943), but since they signify nothing about output they are of no interest in the present context.

Acknowledgement is due to Professor G. Humphrey for permission to work in the laboratory of the Institute of Experimental Psychology. The writer is grateful to his supervisor, Mr. J. A. Deutsch, for encouragement and critical discussion at every stage of the work.

REFERENCES

- APPLEZWEIG, M. H. (1951). Response potential as a function of effort. *J. comp. Psychol.*, **44**, 225-235.
- HULL, C. L. (1943). *Principles of Behaviour*. New York.
- MAATSCH, J. L., ADELMAN, H. M. and DENNY, M. R. (1954). Effort and resistance to extinction of the bar-pressing response. *J. comp. Psychol.*, **47**, 47-50.
- MONTGOMERY, K. C. (1951). An experimental investigation of reactive inhibition and conditioned inhibition. *J. exp. Psychol.*, **41**, 39-51.
- MOWRER, O. H. and JONES, HELEN, M. (1943). Extinction and behavior variability as functions of effortfulness of task. *J. exp. Psychol.*, **33**, 369-385.
- OSGOOD, C. (1953). *Method and Theory in Experimental Psychology*. New York.
- SKINNER, B. F. (1938). *The Behaviour of Organisms*. New York.
- SOLOMON, R. L. (1948a). The influence of work on behavior. *Psychol. Bull.*, **45**, 1-37.
- SOLOMON, R. L. (1948b). Effort and extinction rate: a confirmation. *J. comp. Psychol.*, **41**, 93-101.
- STANLEY, W. C. and AAMODT, MARJORIE S. (1954). Force of responding during extinction as a function of force requirement during conditioning. *J. comp. Psychol.*, **47**, 462-464.
- THOMPSON, M. E. (1944). Learning as a function of the absolute and relative amounts of work. *J. exp. Psychol.*, **34**, 506-515.

DIRECTIONAL INFORMATION AND APPARENT MOVEMENT*

BY

MALCOLM A. JEEVES and JEROME S. BRUNER

From the Psychological Laboratory, University of Cambridge and Department of Social Relations, Harvard University

The number of alternative directions that a point of light can take in the apparent movement produced by the successive exposure of two points of light may be considered its directional information. The present experiment is concerned with the relationship between such directional information and the threshold between apparent movement and apparent successiveness, the original hypothesis being that the greater the directional information, the higher the threshold of movement. In fact, contrary results were obtained. Thresholds were obtained under conditions where movement could occur in either eight directions or only in two directions, better and more persistent movement being found under the former conditions. But this result is obtained only when one measures thresholds by an ascending method: increasing the time separation between light points from optimal movement to successiveness. An hypothesis is proposed to explain the findings stated in terms of "attentional disarticulation."

Conventionally, studies in apparent movement are conducted somewhat as follows. Two stimuli are used (lines or dots or points of light) and these are presented successively, separated in time and space from each other. If the duration and intensity of the two stimuli presented is appropriate, if the spatial and temporal separation between them is appropriate, and if the instructions and context are appropriate, apparent movement will be seen. That is to say, the first stimulus will be seen to "move over" to the position of the second stimulus. The effect is quite compelling and the "moving" neon signs with which the streets of our cities are now blighted is sufficient evidence that it does not require delicate laboratory conditions for its evocation. Graham (1951) has recently restated Korte's laws (1915) of apparent movement relating the intensity and duration of the stimuli used and the spatial and temporal separation between them. And Neuhaus (1930) has shown that instructions and practice can alter apparent movement thresholds while Jones and Bruner (1954) have presented data indicating that "optimal" movement can be more readily produced, and even the direction of movement influenced by presenting the movement stimuli in contexts where the probability of movement normally occurring is greater.

None of the studies thus far reported in the literature has taken into account what appears to us to be a very obvious variable: the directional information contained in the movement perceived. The variable is easily described. As in the conventional experiment, let us use two stimuli, s_1 and s_2 , the former being presented first in time. The positions of s_1 and s_2 may always be the same, in which case the task of the subject is simply to report whether he perceived movement, two successive stimuli, or two simultaneous stimuli. Or the position of the second stimulus with respect to the first may be varied from presentation to presentation. That is to say, for any given subject, the number of alternative positions of s_2 with respect to s_1 may be altered. s_2 may appear to the north or south of s_1 , or it may appear to the north,

* The present research was carried out as part of the Cognition Project at Harvard University.

south, east, or west. The greater the number of alternative directions that movement can take in a given situation, the greater the directional information the movement may be said to contain (using information, of course, in the sense in which Shannon has proposed).

It was our hypothesis at the outset of the present study that the larger the number of alternative directions that apparent movement could take, the poorer would the apparent movement be—i.e., the higher would be the threshold for apparent movement. The argument seemed simple. If movement could occur in one and only one direction, then the subject would be "set" for it and somehow "set" would serve to facilitate movement much as it had done in the Jones and Bruner (1954) study where constraining context facilitated movement. Thus, if one held the intensity and duration of the two stimuli constant and maintained a constant spatial separation between stimuli, varying only the pause between stimuli, then one should find that the threshold between simultaneity and movement should be at shorter time intervals, and the threshold between successiveness and movement should be at longer time intervals as one reduces the amount of directional information carried by the movement.

This seemed a very reasonable hypothesis and it is one that all of our colleagues predicted when asked to do so, and those interested in information theory showed the most conviction of all on the point. The fact of the matter is that results of quite contrary nature have been obtained and perhaps it is just as well to turn directly to the experiment before commenting further on our hypothesis that proved so contrary. Only one point need be made before doing so: The information model may not be so directly relevant to the case we are considering as it first appeared. If between two points S_1 and S_2 , exposed successively, one could increase the number of alternative intervening states possible from three (simultaneity, movement, or successiveness) to some larger number, then one might reasonably expect a change in the threshold of apparent movement according to information theory. Increasing *direction* of movement need not, on strictly logical grounds, be expected to have this effect. One might argue, indeed, that the present experiment might more profitably be considered as an instance of the study of the distribution of attention: what happens to the threshold of apparent movement when expectancy is limited with respect to its direction. These are matters better treated after considering the experiment in proper detail.

PROCEDURE

Sixteen Harvard students were used as subjects. The stimuli for apparent movement were two pinpoint spots of fluorescent white light appearing in a black unstructured field in a Harvard-Gerbrands tachistoscope. The two pinpoints of light were 0.5 in. apart and were 22 in. from the eyepiece of the subject. The lighting used in each half of the tachistoscope was two 75-watt, 250-volt fluorescent tubes placed behind an aperture approximately 0.5 millimeter in diameter. The spots of light were seen in a visual field either through or on a half-silvered mirror, the centre point of which was 4-13/16 in. from the centre of the eyepiece. The half-silvered mirror had equal transmission and reflectance characteristics so that the two points of light were of equal apparent intensity.

In the presentation of the movement stimuli the first light appeared in the centre of the black field for 1.0 second—a length of exposure that permitted subjects to use this first light as a fixation point. The second point of light appeared for 0.35 second and the time interval between the appearance of the first and second lights was varied for the determination of movement thresholds in a manner presently to be related. The subjects wore headphones during the experiment and each time the lights were about to appear in the tachistoscope white noise was transmitted through the phones with sufficient loudness to mask all outside noises which might either distract the subjects or give them any cues from the timing mechanism. Thus, the white noise served both as a warning signal and a masking noise.

The instructions given to the subjects were designed to indicate the directions in which they should expect the second pinpoint of light to appear in relation to the first. Under one set of conditions the subjects were told that the second spot would appear in any one of *two* possible directions and *only* in these two directions. Thus for example one subject would be told to expect movement only to the North or South (N-S) of the first point of light, another in the E-W direction and so on. The four pairs of directions used were N-S, E-W, NE-SW, and NW-SE, which were allotted randomly to the subjects.

Under the second set of conditions subjects were told that they could expect the position of the second spot of light to be in relation to the first spot of light at any of the *eight* points of the compass: E, NE, N, NW, W, SW, S, SE.

Four threshold determinations were made on each subject, two going from apparent successiveness to apparent movement and two from apparent movement to apparent successiveness (i.e. two ascending and two descending series). The method of limits was employed. For the ascending series, the first presentation was given at a temporal separation of 50 milliseconds between the two stimuli, and on successive presentations the separation was increased progressively by 20 milliseconds until the subject gave three successive reports of "no movement." The starting point of the determination, 50 milliseconds, produces excellent movement at the spatial separation used. For the descending series, the first presentation of the two light points was at a temporal separation of 600 milliseconds (a duration that for most subjects produces the impression of two successive light points) and it was reduced on successive exposures by 20 milliseconds until movement was finally reported three presentations running.

It would be sensible to consider the ascending series a determination of the *threshold of apparent successiveness*, the descending series a determination of the *threshold of apparent movement*.

Our subjects were divided into four subgroups each given somewhat different orders of testing. These are as follows, and subjects were assigned to the subgroups randomly.

I	II	III	IV
"2" ascending	"2" descending	"8" ascending	"8" descending
"8" ascending	"8" descending	"2" ascending	"2" descending
"2" descending	"2" ascending	"8" descending	"8" ascending
"8" descending	"8" ascending	"2" descending	"2" ascending

By the use of such a counterbalanced design we are able to determine the effect of the 2- and 8-directional expectancies, the method of computing threshold, the effects of practice or adaptation as one goes from the first to the second half of the experiment, and finally the effect of having a 2-directional or 8-directional expectancy initially.

The subjects were instructed as follows:—

"This is an experiment designed to investigate certain characteristics of apparent visual movement. I want you to look into the eyepiece of this tachistoscope and at intervals I shall show you two points of white light appearing in the black field. Under certain conditions like this (subject shown apparent movement condition until he gets good appearance of apparent movement) you will appear to see not two separate spots, but instead one spot of light appearing to move across the field. Under another set of conditions you will see the successive appearances of two separate and independent spots of white light like this (subject shown successive condition). The second spot of light will appear either vertically upwards or vertically downwards [or whichever condition is being used] and your task is to report to me after each exposure the direction of the second spot and whether it was a case of apparent movement or successive appearances. There may be some instances when you are doubtful; in that case I want you to say so and to report which of the two conditions it most resembles.

I also want you to wear these headphones—they serve two purposes: firstly, to cut out all extraneous noises so that you will not be distracted during the experiment and secondly, to give you warning to be ready for the lights to appear. Each time before the lights appear I shall turn on the white noise like this and then turn it off again as soon as the lights have disappeared. You should relax except when the noise is on and also rest your eyes between exposures. Is that clear? Do you have any questions?"

RESULTS

A summary of the results is contained in Table I: the amount of temporal separation in milliseconds between two successively presented light points to produce the appearance of visual movement. The longer the separation, since we are working at the border between apparent movement and apparent successiveness, the more sensitive or "lower" the movement threshold. That is to say, if a subject reports apparent movement at 300 milliseconds of separation between the light points and reports apparent successiveness at 320 milliseconds, his threshold is lower than one who reports movement at 220 milliseconds and successiveness at 240. Thus, the higher the temporal value, the lower the threshold (in the conventional sense) for apparent movement.

TABLE I

THRESHOLD OF APPARENT MOVEMENT FOR VARIOUS CONDITIONS EXPRESSED IN MILLISECONDS OF TEMPORAL SEPARATION BETWEEN TWO STANDARD LIGHT POINTS

	<i>2 directions expected</i>	<i>8 directions expected</i>	<i>Combined</i>
All conditions	191	226	208
Ascending method	213	268	240
Descending method	169	184	178
First half determinations	176	231	203
Second half determinations	208	220	214
2 directions initially	206	245	226
8 directions initially	176	208	192
Ascending: first half	231	310	270
Ascending: second half	195	225	210
Descending: first half	120	152	136
Descending: second half	218	215	216
Ascending: 2 initially	210	272	241
Ascending: 8 initially	218	262	240
Descending: 2 initially	202	215	209
Descending: 8 initially	160	152	156

Several findings contained in the table want restating. The first and most obvious is that there is a significant increase in sensitivity to apparent movement when one operates with an 8-directional expectancy in contrast to a 2-directional one. In short, the greater the amount of directional information carried by the movement, the lower the threshold of movement appears to be. The gross difference observed is statistically reliable as determined either by the use of a *t*-test on difference scores for individuals (*p* is less than 0.001) or by a grosser analysis of variance on unpaired data (*p* is less than 0.05).

A further inspection of Table I shows that there is a marked difference in the results obtained by the use of ascending and descending methods—the former measuring, as we have suggested, a threshold of apparent successiveness; the latter,

a threshold of apparent movement. For example, if one disregards the kind of expectancy subjects work under, one finds that the threshold obtained by the ascending method is 240 milliseconds, by the descending method 178—a difference of more than 60 milliseconds. It is not surprising then that the statistical reliability of the difference between the two methods is very high. It is the conventional practice, in dealing with psychophysical threshold data, to combine the results obtained from the two procedures used in the method of limits. But there is good reason not to do so in the present case—good reason that goes beyond the discovery that method of determination is a significant source of variance in our data. Note that when one uses the method of determining the threshold of successiveness (the ascending procedure), one obtains a marked difference between 2- and 8-directional expectancies. Movement persists for an 8-directional expectancy at a temporal separation 65 milliseconds greater than for a 2-directional expectancy. This difference is all but washed out when one uses the descending method, based on first reports of apparent movement. In view of these facts, we have decided to treat the two sets of data separately from a statistical point of view—concretely to perform separate analyses of variance on the two sets of data, the ascending and descending.

As for the analysis of the descending data, none of the sources of variance studied show a significant effect. Neither the kind of expectancy under which the subject operated, nor the order in which determinations were made, nor the difference between the first and last two threshold determinations yields a significant effect. This is partly because general variability is greater under these conditions, and partly because the systematic effects introduced by the various conditions studied was less. In short, the data obtained by measuring the threshold of movement proper (in contrast to measuring the threshold of successiveness) are inconclusive with respect to our principal hypothesis—the effect of directional information—and with respect to any other relevant variables studied.

With respect to results obtained by the ascending method—going from apparent movement to apparent succession—the differences obtained are quite clear cut. Directional expectation is a significant source as we have noted (at a confidence level of 0.05), with a very markedly greater sensitivity to movement characterising the 8-directional expectancy. There is also a notable tendency for the threshold to become less sensitive in the second half of the experiment, regardless of type of expectancy ($p = 0.05$). The order in which determinations are taken—whether a 2- or 8-directional expectancy is given first or whether ascending or descending method is used first—appears to make no difference and contributes virtually nothing to variance. Finally, there is a striking tendency for sensitivity to movement to decrease generally during the second half of the experiment if one has started with a 2-directional expectancy, whereas there is an equally marked tendency for sensitivity to movement to increase if one has begun with an 8-directional set ($p = 0.01$). These findings are without regard to whether, after the first determination, one is operating with an 8- or 2-directional set.

One other general result should be mentioned, although analysis of variance shows that it is not significant. This tendency is for the difference between the 2- and 8-directional movement to wash out over the four threshold trials. That is to say, thresholds begin further apart than they end up—a result to be noted in the fourth and fifth row of Table I. This is a finding based on combining the data for ascending and descending methods and should, therefore, be taken with a grain of salt. It may be, however, that there is a regression effect working against the differences we have been discussing. No such effect can be found in an inspection of the ascending and descending data treated separately.

DISCUSSION

The first point that wants making is that one does well in measuring thresholds of "apparent movement" to make a distinction between the threshold obtained by studying when movement *ceases* and the one obtained by studying when it *starts*. The ancient tradition, when one uses the method of limits, is to combine the results of the ascending and the descending measures. In the present study, the two methods yield quite different results and one combines them at one's own peril. The fact of the matter is that several conditions that affect the threshold of successiveness have been found *not* to affect the threshold of movement.

Let us reconstruct the two situations in which testing is carried out. In the one—the determination of a threshold of successiveness—a subject is first presented with two stimuli separated in time in such a manner as to produce good apparent movement. On subsequent exposures, the time gap between the two stimuli is lengthened until, finally, a point is reached where the subject tells you that he no longer sees movement. Under such conditions, we find that it makes a considerable difference whether a subject is set to see movement taking place in any of eight directions or only in two directions. When one operates, on the other hand, by starting the subject with stimuli separated in such a way that they appear *not* to be moving but to be succeeding each other in time, and then gradually *decreases* the time gap until movement is produced, no such effect is found. What produces this effect under the first condition, and why does it not appear in the second?

We believe that the results can, in general terms, be discussed in terms of the "attentional disarticulation" of the perceptual field under instructions to expect movement in any of eight directions (as compared with either of two directions). Why in the one testing condition rather than the other? When the subject is set to expect movement in either two or eight directions and is then shown a stimulus field where movement is occurring (the ascending condition), the expectancy is appropriate to the conditions present in the field before him. When the field contains *no* movement, the set is inappropriate. Moreover, the field in which movement is occurring is, if we are to take seriously the various neural theories of apparent movement, in a state of disequilibrium in which "short-circuiting" or "communication" between adjacent brain fields is operative. This is not the case when the points of light are so separated at the outset of measurement as to produce the appearance of two stationary light points succeeding each other.

We may next ask, "But why does an 8-directional set produce greater sensitivity to movement, and what is meant by 'attentional disarticulation' under these conditions?" Consider the matter in more detail.

If one distributes one's attention over eight possible pathways of apparent movement, we would hold, this may be tantamount to reducing the kinds of constraints that operate in a field to keep events and points in that field segregated from each other. Under these circumstances, "short circuiting" or whatever it is that makes successively presented lights appear like a single light moving from one point to another should become more likely. This would certainly not be inconsistent with other observations on the distribution of attention as a disarticulating factor in perceptual organisation. Certainly Newhall's data (1923) would support the view: he found the more focalized attention was by a foresignal, the greater the ability of subjects to segregate and discriminate weak tactile stimuli. And the data on more evolved cognitive operations would certainly indicate that reduction in capacity to structure tasks occurs when attention is distributed (e.g. Sterzinger, 1924). But perhaps most striking of all are the results of experiments on the information receiving

capacity of human operators. For example, Miller, Heise, and Lichten (1951) have shown that at a constant signal-to noise ratio, the ability to recognize orally presented words declines as the number of alternative words in the test vocabulary is increased. It has been conventional to interpret such findings in terms solely of the coding problem involved in matching a certain number of inputs with a certain number of coding categories, the general assumption being that the larger the array, the greater the number of binary decisions to be made in such a sorting of inputs among coding categories. But it may well be too, that the distribution of expectancy (or attention) over a wider array of alternatives may serve to cut down the degree of sensory organisation or articulation present in the perceptual field.

We do not, as the reader can see, have a very satisfactory explanation of the findings we have reported. It had been our expectation at the outset that quite opposite results *would be obtained*, and these might well have been handled in conventional informational terms. The best we are able to do is to propose an explanation in terms of the distribution of attention and its effects. However unsatisfactory the explanation, one point is quite clear. Directional information is a contributing variable in the array of factors that determine thresholds of apparent movement. It is one in need of closer examination than we have been able to give it in these pages. It is our hope that this preliminary report on its effects may lead to a fruitful revisit to the phenomena of apparent movement in general—phenomena whose study may throw light on more aspects of perceptual organisation than suggested by the polemical questions with which Wertheimer awakened interest in the topic two generations ago.

REFERENCES

- GRAHAM, C. H. (1951). Chapter 23, Visual Perception. In *Handbook of Experimental Psychology*. Ed. S. S. Stevens. New York.
- JONES, E. E., and BRUNER, J. S. (1954). Expectancy in apparent visual movement. *Brit. J. Psychol.*, **45**, 235-240.
- KORTE, A. (1915). Kinematoskopische Untersuchungen. *Z. Psychol.*, **72**, 193-296.
- MILLER, G. A., HEISE, G. A., and LICHTEN, W. (1951). The intelligibility of speech as a function of the context of the test materials. *J. exp. Psychol.*, **41**, 329-335.
- NEUHAUS, W. (1930). *Experimentelle Untersuchung der Scheinbewegung*. *Arch. ges. Psychol.*, **75**, 315-458.
- NEWHALL, S. M. (1923). Effects of attention on the intensity of cutaneous pressure and on visual brightness. *Arch. Psychol.*, **9**, 5-75.
- STERZINGER, O. (1924). Zur Prüfung und Untersuchung der abstrakten Aufmerksamkeit. *Z. angew. Psychol.*, **23**, 121-161.

THE INFLUENCE OF AGE, SEX, AND SPECIALIZED TRAINING ON THE VERTICAL-HORIZONTAL ILLUSION

BY

P. FRAISSE and P. VAUTREY

*From the Laboratory of Experimental and Comparative Psychology
of the Sorbonne, Paris*

With the aim of investigating the perceptual processes responsible for perceptual illusions, we have studied the influence of age, sex and specialized training on two forms of the vertical-horizontal illusion. In Figure A, the vertical is a short distance away from the horizontal; in Figure B, the vertical meets the horizontal segment at its mid-point. Both figures were presented in brief (0.2 second) and long (1.0 second) tachistoscopic exposure, and in unlimited exposure.

When time is unlimited, Illusion A increases up to 9-10 years and then remains constant and Illusion B decreases as a function of age and chiefly as a function of the degree of scientific training. The illusions are stronger in tachistoscopic perception than under conditions of unlimited exposure. Both illusions also increase in tachistoscopic perception with age except for science students. Under condition of unlimited time, women have stronger illusions than men. This difference does not occur in tachistoscopic perception.

These results may be interpreted in terms of the interaction of three processes: interaction among the elements of a figure (field effects); perceptual activity which makes possible the passing from syncretism or multiplicity to organized perception; and the activity of measuring the elements relatively to each other. This latter activity develops with cultural level and especially with scientific training. It is more effective with men than with women, due, it is suggested, to the spatial factor being higher in men.

INTRODUCTION

Binet (1895) as early as 1895 showed the usefulness of visual illusions for studying the development of perceptual processes by demonstrating that illusions like the Müller-Lyer decrease with age, while others, like the optico-kinesthetic illusion or Demoor's weight illusion, increase with age. This field was subsequently little explored until the important work of Piaget once again introduced the problem.

Wursten (1947), one of Piaget's students, reported that, in addition to the two types of illusions pointed out by Binet, there was a third type represented by one form of the vertical-horizontal illusion which increase up to the age of 9-10 years and decrease thereafter. Among the latter, he studied the illusion in which the vertical is over-estimated when shifted (some distance from the horizontal) relatively to the horizontal. This fact, if verified, would lead to a theory of perceptual development rather different from that called for by Binet's results. Binet's experiments indeed suggested that some illusions were of a primitive nature, since they decreased with age, while others were due to a learning factor. Without first entering into a theoretical discussion from the start, we wondered whether Wursten's data were not due to an artifact. This question led us to study the vertical-horizontal illusion as a function of the duration of perception and individual differences related to age, sex, and educational background.

METHOD

Material. Two figures for the vertical-horizontal illusion (over-estimating the vertical with respect to the horizontal) were used. The first, figure A, was like an L, the vertical segment being 3 cm. long and displaced 5 cm. to the left of the horizontal, which was also

3 cm. long. In figure B, the two segments (also 3 cm. long) formed an inverted T. In the first figure, the distance between the two segments made comparison difficult; in the second, the illusion was increased by the under-estimation of the horizontal segment, which was cut at its mid-point.

These figures were presented under three different temporal conditions: (a) brief exposure (0.2 second), the equivalent of a very rapid glance; (b) average speed (1.0 second), allowing for a rapid exploration; (c) unlimited time, a condition allowing not only for exploration but also for an analysis of the figure, comparisons, and constructs. A tachistoscope was used for the first two conditions. It consisted of a projection-lantern which projected slides on a ground-glass screen with its fixation point at the geometrical centre of the figures. The subjects were seated 1 m. from the screen. Under the condition of unlimited exposure, the figures were presented on light cardboard cards, about 50 cm. from the subject. The dimensions of the figures and the area of the background were the same for both the screen and the cards. The constant method was used throughout the experiment. We determined upper and lower thresholds by utilizing five series of stimuli. The mean of these thresholds was taken as the point of subjective equality. From this basis, it was easy to compute the systematic error attributable to the illusion.

Subjects. Five groups of subjects were studied: (I) 36 six-year-old children (5 years, 9 months–6 years, 4 months) taken from the public nursery schools of Paris, 7 boys and 9 girls for the tachistoscopic trials, and 10 children of each sex for the trials under the condition of unlimited time; (II) 55 9–10-year-old children taken from the primary schools of Paris, 9 children of each sex for the tachistoscopic task, and, under the condition of unlimited time, 17 boys and 20 girls; (III) 47 adults of low educational level (with less than high school education)—12 men and 15 women, from a centre for accelerated professional training, in the tachistoscopic tests, and 6 men and 14 women, parents of the primary school children, in the tests under the condition of unlimited time; (IV) 40 graduate students in liberal arts, 11 men and 8 women with the tachistoscope, 11 men and 10 women under the condition of unlimited time; (V) 43 graduate students in science, 10 men and 10 women with the tachistoscope, 11 men and 12 women under the condition of unlimited time.

We used three groups of adults because it did not seem legitimate to compare children with just any group of adults, unless it was representative of the entire population. The 6- and 9-year-old children were quite homogeneous with respect to background, but the adult groups differed with respect to education and training after adolescence.

In the unlimited-time test, Figure B was always presented after Figure A. For the tachistoscopic tests, we devised an experimental design in which the order of presentation of the figures and the exposure-time order were counterbalanced between subjects. Four series of measures were determined for each subject. These four series were divided into two sessions with the adults and in four sessions with the youngest children so as to avoid the effects of fatigue.

RESULTS

The overall results are presented in Tables I and II. In general, the results show the following:—

- (1) Over-estimation of the vertical occurs to the same degree at every age level. Further differences which we shall study are minor in comparison with the fundamental fact.
- (2) The illusion is about twice as great with Figure B (the inverted T) than with Figure A. The proximity of the segments being compared and the bisection of the horizontal by the vertical in the case of the inverted T might on first analysis account for this result.
- (3) At all age levels and for both figures, the illusion is stronger with the tachistoscopic exposure than under the condition of unlimited time (the same fact was brought to light by Sleight and Austin (1952)) and slightly stronger with short than with long tachistoscopic exposure at each level. The less opportunity a subject has to explore the figure (the more global his perception of the field), the stronger is the illusion.

TABLE I

FIGURE A (| —)

MAGNITUDE OF THE SYSTEMATIC ERROR (in mm.)

Exposure time	Nature of the groups	6-yr. (N)	9-10-yr. (N)	Low educ. level adults (N)	Arts students (N)	Science students (N)
0.2 sec.	M	3.5 (7)	4.0 (9)	3.8 (12)	5.1 (11)	4.0 (10)
	F	4.4 (9)	4.6 (9)	4.4 (15)	4.7 (8)	4.1 (10)
	Mean	3.9	4.3	4.1	4.9	4.0
1 sec.	M	3.6 (7)	3.4 (9)	3.7 (12)	4.8 (11)	3.3 (10)
	F	4.0 (9)	4.3 (9)	3.7 (15)	4.5 (8)	3.6 (10)
	Mean	3.8	3.8	3.7	4.6	3.5
Unlimited time	M	2.5 (10)	3.2 (17)	3.4 (6)	3.0 (11)	3.4 (11)
	F	2.6 (10)	4.3 (20)	4.0 (14)	4.3 (10)	3.7 (12)
	Mean	2.5	3.8	3.7	3.6	3.6

Significant differences

Unlimited time { Between sexes:
at 0.05 level at 9-10 years and among liberal arts students.
Between ages and levels:
at 0.02 level between 6-year and 9-10-year girls.

Tachistoscopic perception: The analysis of variance of the results taking into account the difference of the number of subjects in each group shows:

A significant variation at the 0.10 level among ages and levels.

" " " " " 0.01 " between sexes.

" " " " " 0.01 " between the two exposure times.

All interactions between sources of variance are non-significant.

TABLE II

FIGURE B (⊥)

MAGNITUDE OF THE SYSTEMATIC ERROR (in mm.)

Exposure time	Nature of the groups	6-yr. (N)	9-10-yr. (N)	Low educ. level adults (N)	Arts students (N)	Science students (N)
0.2 sec.	M	5.8 (7)	9.6 (9)	9.9 (12)	10.4 (11)	8.7 (10)
	F	7.5 (9)	7.4 (9)	8.9 (15)	10.4 (9)	7.0 (10)
	Mean	6.7	8.5	9.5	10.4	7.9
1 sec.	M	5.8 (7)	9.7 (9)	9.2 (12)	9.0 (11)	7.6 (10)
	F	7.3 (9)	7.5 (9)	7.8 (15)	9.1 (9)	7.5 (10)
	Mean	6.5	8.6	8.5	9.1	7.5
Unlimited time	M	6.5 (10)	6.0 (17)	4.5 (6)	5.3 (11)	1.9 (11)
	F	6.4 (10)	6.6 (20)	6.6 (14)	6.8 (10)	5.1 (12)
	Mean	6.5	6.3	5.6	6.0	3.5

Significant differences

Unlimited time: Between sexes at 0.01 level among science students.

Between ages and levels among male subjects, significant at 0.01 level between science students and all other groups.

Tachistoscopic perception: The analysis of variance of the results taking into account the number of subjects in each groups shows:

A significant variation at the 0.01 level among ages and levels non-significant variation between sexes.

A significant variation at the 0.01 level between the two exposure times.

The following interactions: (a) age-level $\times$ sex; (b) sex $\times$ exposure-time; (c) age-level $\times$ exposure time $\times$ sex, are significant at the 0.01 level.

The Influence of Age

With Figure A, the illusion increases from 6 to 10 years under condition of unlimited exposure while it does not increase under condition of tachistoscopic perception. With Figure B, the opposite trend appears with an increase in the case of tachistoscopic perception, but no change or a slight decrease in the case of unlimited exposure. Since the same children were not tested under both conditions of unlimited and tachistoscopic exposure, interpretation of results is hazardous. We may note, however, that Wursten also found the effect with Figure B to increase over the same age range. For the six-year-old child, Figure A seems to be little structured because of the distance between the two segments, and the illusion apparently is weakened consequently. Structuration seems to increase with age. In our opinion, the fact that the two lines do not meet plays a major rôle. While the Müller-Lyer illusion decreases slightly with age, Triche and Hartmann (1933) found that the illusion, on the contrary, increases with age, when in the figures presented, the extremities only are used as shown here: [] []

We likewise found that the perception of unfilled time (marked off by two separate sounds) was very difficult for young children, while that of filled time (continuous sound) was easy and precise (Fraisie, 1948).

At six years of age, the illusion seems distinctly stronger with tachistoscopic than with unlimited exposure. Does the impossibility of viewing the two lines separately increase the illusion? This hypothesis is indirectly confirmed by the fact that the Figure B-illusion increases markedly in tachistoscopic perception between the ages of six and nine, while the known fact that this illusion decreases slightly in unlimited exposure is also confirmed. Brief exposure which increases the field-effect by inducing perception of the whole must already play an important rôle at the age of six. This fact should be verified with other illusions which do not bring into play the phenomenon of direction but solely that of size, like the Müller-Lyer illusion.

Why does the illusion of the inverted T decrease slightly with age, as does the Müller-Lyer illusion, under condition of unlimited exposure? The classical interpretation postulates an analytical attitude which breaks down the field and minimizes the interaction of the segments. According to Piaget, it is perceptual activity which operates. These factors certainly play a rôle, but subjective reports of subjects show that the older the subjects are, the more they try to make comparisons and constructs in an effort to *measure* one line by the other. The measurements consciously made and related to intelligent activity, make some correcting possible taking due account of the inadequacy of the means used (perceptual and imaginative activity). This interpretation is supported by a comparison of the results obtained from various groups of adults having different intellectual training.

The Influence of Educational Background

Let us first consider the results obtained under the condition of unlimited exposure. In general, we observe the fact that educational background does not affect the results with Figure A. All our adult groups yield results similar to those of the

9- to 10-year-old children. With Figure B (inverted T), however, there is a marked effect of educational background, particularly in our group of science graduates. Actually, this result can be explained, as we suggested in the above paragraph, by the comparison of the two segments in a genuine operation of measurement.

This fact was verified by comparing the magnitude of the systematic error with the percentage of subjects who consciously attempted to build constructs in an interview following the test.

This verification was carried out not only with the three groups of women already mentioned, but with two additional female groups, one having an elementary education only, and the other consisting of psychology students.

The five groups took the unlimited exposure tests for the two illusions and Table III shows the relationship between the magnitude of the illusion and the percentage of subjects consciously attempting constructs.

TABLE III

GROUPS	FIGURE A		FIGURE B	
	<i>Systematic error</i>	<i>Conscious constructs</i>	<i>Systematic error</i>	<i>Conscious constructs</i>
Elementary education only N = 19	4.2 mm.	11%	8.7 mm.	47%
Liberal arts graduates .. N = 10	4.3 mm.	7%	6.8 mm.	60%
Less than high school .. N = 14	4.0 mm.	10%	6.6 mm.	64%
Science graduates .. N = 12	3.7 mm.	33%	5.1 mm.	100%
Psychology students .. N = 20	2.9 mm.	50%	4.6 mm.	80%

For both illusions, we see that the systematic error decreases as the percentage of conscious attempts at constructs increases.

However, the constructs are much less frequent with Figure A than with Figure B (inverted T). But in Figure A, the distance between the two segments makes an imaginative measuring operation very difficult, while with the inverted T, it is just the reverse, the bisection of one segment by the other making such measurement easier. This difference explains why one of the illusions decreases with the level of scientific and technical training while the other does not. In the case of Figure A, one group, however, has a weaker illusion than the others. This is the group of psychology students who, owing to past experience, succeed in increasing the percentage of constructs and in somehow reducing the magnitude of the illusion.

This result, which clarifies the process whereby some illusions decrease with age, also affords a new interpretation of some earlier results. Crosland, Taylor and Newson (1929) for example, found that there was no correlation between intelligence level and susceptibility to the Müller-Lyer illusion. But intelligence level was measured by a general intelligence test. Our results show that the essential factor is training which is more or less in pre-established harmony with qualitative differences in intelligence. The arts students showed test results very close to those given by adults having a low educational level, while psychology students, who are certainly no more intelligent than the art graduates, were the least susceptible to both illusions.

On his part, Wursten has shown, on the basis of figures of the same type as our Figure A, that the illusion increases with age up to about 9-10 years (which our results confirm) then decreases slightly. Piaget encountered the same development in many other investigations in the field of perception. This type of illusion—at least in

Wursten's research—may be explained by the fact that the groups of older children and adults are superior both socially and intellectually, to the rest of the population. After 10–12 years, a selection process operates in the schools. The lack of homogeneity of the training background of the groups studied might then explain a bias in the results. Furthermore, Wursten verbally reported to us (in a personal communication) that his groups of adults were composed of cultured people with a preponderance of psychologists. We think that Wursten would have found the same inflection point at the age of 9–10 had he used the Demoor illusion with the same subjects. This hypothesis remains to be verified.

But the influence of training does not only enable the subject to make better use of conscious constructs. In the tachistoscopic tests, we also find differences among the adult groups, and these differences tend in the same direction as in the unlimited exposure tests, but to a lesser degree. Let us consider, for example, the greatest difference: the science graduates have a weaker illusion than the liberal arts graduates at the speed of 0.2 second and at 1 second. This fact suggests the hypothesis that specialized training has caused the science students to develop an analytical attitude which manifests itself even in a rapid glance.

C. *The Influence of Sex*

Under conditions of tachistoscopic exposure there is practically no systematic difference between boys and girls. Under condition of unlimited exposure, on the other hand, the boys from the age of 9–10 years are superior to the girls and this difference, which is always in the same direction, is significant for several groups, with one figure or the other. It is greatest for the groups of science graduates in the case of the inverted T. Although the same proportion of male and female students used conscious constructs, the magnitude of the illusion is 5.1 mm. for the female students and only 1.9 mm. for the male students. This result leads us to take into account a new variable. We know that factorial studies have brought to light a spatial factor (S factor). In tests highly saturated with this factor, boys are superior to girls. This factor (S_2) is interpreted as the aptitude for mental manipulation of spatial figures. We think that our results may be explained as due to this. Under conditions of unlimited exposure but not with tachistoscopic exposure, differences in training have an increased effect, by enabling the aptitude for manipulating spatial forms to give rise to measuring operations.

DISCUSSION AND CONCLUSIONS

If we try to bring together the various aspects of our results, it is possible to distinguish the action of several hypothetical processes the combination of which would account for our experimental results.

(1) *Perceptual organization of the figures.* As the Gestalt school has brought out, the different parts of a figure are interrelated; these interactions are field-effects. These effects which are ontogenetically very primitive can, however, increase with age through a better interrelating of the elements when they are separated as in Figure A, or in the process of passing from syncretism to organization. Indeed, some field effects are perhaps hidden in the case of young children by the syncretic character of their perception; failure to distinguish among the parts might reduce their interaction. This last aspect would, according to us, explain the fact that the illusion in Figure B (inverted T) increases with age in tachistoscopic perception. For in this case the process of analysis, which we are going to describe, is not sufficient to compensate for the field effects themselves. This fact which has been brought to light for illusions of direction should be ascertained for illusions of size.

This increase of the field effects through a better organization of the figure can be explained by the development of perceptual activity. With increasing age, the exploration of the figure becomes more complete.

(2) *Analysis of the figure into its elements.* Perceptual activity can change the direction of its action on one hand as a function of age, on the other, as a function of the different field effects. Its organizing effect can become an analytical effect which, through a relative isolation of the perceptual elements, would diminish the interactions of the field. The Müller-Lyer illusion, for example, diminishes if its measurement is often repeated (as a function of the number of presentations). We have also seen that illusions were always weaker in unlimited perception than in tachistoscopic perception. Perceptual activity can give rise to a true analytical attitude in individuals, particularly as a function of their training. Technicians and scientists may learn how to isolate relatively the constituent parts of a figure at first glance.

The increase or decrease of illusions with age could, in fact, be explained by an identical process: perceptual activity the direction of which can change according to the relations existing among the elements of a perceptual whole. Where the links are weak (or confusion too great) this activity would have an organizing effect. Where they are strong it would have an analytical effect, on which specialized training would exert a definite influence.

(3) *Relative measurement of the elements of the figure.* Beside perceptual activity, our results show the fact that the individual is able to put the various elements of a figure into a quantitative relationship with each other. These measuring operations, which are made by using perceptual reference points are of an intellectual nature, operational in Piaget's terms. We leave here the field of perception; these measuring operations are evident only when the situation calls for a quantitative comparison of elements, as in the case of our experiments. The number of these constructs is a function of the intellectual training of the subjects and the nature of this training. They play no rôle at the age of 6 and but little at the age of 9. But they are very frequent among adults. In addition to the frequency, we must consider the efficiency of these constructs, which are largely dependent on specialized training and on the aptitude for mental manipulation of figures which shows up in the differences of results given by boys and girls.

REFERENCES

- A. BINET (1895). La mesure des illusions visuelles chez les enfants. *Rev. Philos.*, **40**, 11-25.
 H. R. CROSLAND, H. R. TAYLOR, and S. J. NEWSOM (1929). Practice and improvability in the Müller-Lyer illusion in relation to intelligence. *J. gen. Psychol.*, **2**, 290-306.
 P. FRAISSE (1948). Etude comparée de la perception et de l'estimation de la durée chez les enfants et les adultes. *Enfance*, 199-211.
 R. B. SLEIGHT, and T. R. AUSTIN (1952). The vertical-horizontal illusion in plane geometric figures. *J. Psychol.*, **33**, 279-287.
 A. TRICHE and G. W. HARTMAN (1933). Differential susceptibility of children and adults to standard illusions. *J. Genet. Psychol.*, **42**, 493-498.
 H. WURSTEN (1947). L'évolution des comparaisons de grandeur de l'enfant à l'adulte. *Arch. de Psychol.*, **32**, 1-144.

SEXUAL EXHAUSTION AND RECOVERY IN THE MALE RAT*

BY

FRANK A. BEACH and LISBETH JORDAN

From Yale University

SYNOPSIS

Twelve male rats were left with receptive females and allowed to copulate and ejaculate until they reached a criterion of "sexual exhaustion." They were then retested after 1, 3, 6 and 15 days of sexual inactivity. Following these observations males were tested once each day or once every other day and allowed to achieve a single ejaculation.

In the course of a period of unlimited access to the receptive female males usually need approximately 10 intromissions to produce the initial ejaculation, but successive ejaculations are produced by fewer and fewer intromissions. The time to recover from the effects of an ejaculation increases progressively as exhaustion is approached.

Very few animals copulate when tested 24 hours after sexual exhaustion. Considerably more recovery is evident in tests conducted after a 3-day rest, but it is not complete and rats are not capable of achieving as many ejaculations as they tend to achieve after longer periods of inactivity. As measured by ejaculation-frequency, the curve of sexual recovery is negatively accelerated and probably reaches asymptote after 7 to 10 days of rest.

Various other measures in addition to ejaculation-frequency support this conclusion.

Males allowed to ejaculate once each day or every other day are somewhat less responsive than fully rested animals, but do not show any progressive loss in sexual excitability or capacity.

A working hypothesis is proposed to explain most of the findings. It postulates the existence of an *Arousal Mechanism* which is distinct from a *Copulatory Mechanism*. The ways in which these hypothetical mechanisms are affected by sexual performance and sexual rest are discussed.

INTRODUCTION

Theoretical Background

Sex is commonly listed with hunger and thirst as one of the "primary drives" (Miller & Dollard, 1941), but it has been subjected to only limited systematic analysis. Psychologists who employ this particular classification presumably do so because they believe that when it is investigated "sex drive" will be found to behave in more or less the same manner as hunger and thirst. There are several reasons to question such an assumption. In any event, the term "drive" has acquired so many different meanings and developed so many conflicting connotations that it is of questionable value as a scientific concept. We use it in this introduction in conformation with current fashion, but we do not regard the present investigation as a study of sex "drive." It constitutes an examination of certain factors influencing sexual arousal and copulatory performance.

In the first place, hunger and thirst are related to conditions of tissue need. Sex is not. The notion that tension and pressure in the primary or accessory reproductive apparatus is a source of sex drive is challengeable. Male rats deprived of testes, prostate gland and seminal vesicles copulate normally when treated with androgen (Beach, 1947). Females of the same species in which ovaries, tubes, uterus and vagina are congenitally absent execute the mating pattern under the influence of exogenous ovarian hormones (Beach, 1945).

* This study was supported in part by a research grant (M-943) from the National Institutes of Mental Health, Public Health Service. We are indebted to Allan Goldstein for assistance in the statistical analysis of results.

In the second place, sex differs from hunger and thirst in terms of its biological or evolutionary significance. The latter two can appropriately be designated as drives essential to the preservation of the individual. Sex is logically classified as a drive essential to the preservation of the species. In this respect sex is more closely related to maternal drive. Failure to react appropriately to hunger or thirst can have fatal consequences for the organism. Absence of sex drive, or failure to react to it when it is present does not impair the well-being of the individual, although if such a condition were widespread the existence of the species would be imperilled.

Inasmuch as these two types of drive have quite different biological functions, they may well have quite different evolutionary histories. At any rate there is no justification for the *a priori* assumption that sex drive obeys the same laws and has the same characteristics as hunger and thirst. It does not involve a state of physiological need or metabolic deficit. Its "satiation," or better, its exhaustion, is a catabolic rather than an anabolic process, resting upon the expenditure of energy with no resultant "return" to the individual. Satiation of hunger or thirst restores the organism to an optimal condition. The end-point of sexual exhaustion is fatigue.

A period of deprivation following satiation of the hunger or thirst drives is marked by gradually increasing tissue needs which, if not relieved, eventually result in death. The sequel to sexual exhaustion is a period of recovery, and during a protracted term of deprivation the strength of sexual motivation or capacity first rises and then probably stabilizes at some point near the maximal level. There is some reason to believe that if deprivation is continued for a long enough time sexual responsiveness tends to decrease, but this is speculative.

Earlier Studies and Purpose of the Present Experiment

Previous experiments. As indicated above, systematic studies of sex drive in normal animals have not been numerous. In particular the nature of sexual exhaustion and subsequent recovery of sexual capacity have received very little attention. Casual observations and naturalistic descriptions make it clear that pronounced species differences exist. But even for the most frequently studied animal—the domestic rat, the evidence is incomplete and inconclusive.

The most widely cited experiment of the male rat is that of Warner (1927), who used the Columbia Obstruction Box to measure drive strength. Warner's conclusions were as follows:

The male sex drive appeared to be at its lowest point immediately after a two-hour period during which mating had taken place freely. Recovery was rapid during the first six hours and almost as rapid during the succeeding six. By 24 hours the tendency of males to cross the grid to the female had reached its high point. During the following 6 days there was apparently little change, but what change there was was in the direction of a reduction in the strength of this tendency (p. 159).

More recent investigations lead us to conclude that Warner was not measuring sex drive at all, or at least that observation of *copulatory performance* would have shown that recovery is much slower than his conclusions suggest. For example, Schwartz (1955) trained male rats to press a bar in order to obtain access to a receptive female. Rate of pressing on a variable interval schedule was taken as one measure of sexual responsiveness. An independent measure was the copulatory behaviour. Males were sexually exhausted in the apparatus and retested 1, 3 or 6 days later. Consummatory measures showed that the curve of recovery is negatively accelerated

and approaches asymptote approximately 6 days after satiation. Rate of bar-pressing proved an insensitive measure of sexual recovery, but was significantly correlated with changes taking place in the consummatory response during the course of exhaustion.

Purpose. The purpose of the present investigation was twofold: (1) to measure the course of sexual exhaustion in fully-rested male rats, and (2) to measure the degree of sexual recovery at different intervals after exhaustion. The study was conceived as a normative one which would provide empirical information essential to the planning of subsequent experiments.

SUBJECTS, APPARATUS AND PROCEDURE

Selection and Maintenance of Subjects

Experimental subjects were 12 male rats approximately 150 days old at the beginning of experimentation. Both hooded and albino types were included. All males had been born and reared in our laboratory. Sexually active males were selected on the basis of their mating performance in three or four preliminary tests with receptive females. All animals copulated in the preliminary tests, and 10 males ejaculated. Males were housed individually and were maintained in the experimental room in which the light-dark cycle was controlled by an electric clock. Drinking water and Purina chow were available *ad libitum*, and escarole was added twice weekly.

Apparatus

The apparatus included special observation cages and a recording device. The cages were semicircular, having a straight, glass front, sheet-metal sides and back, and a plexiglass top. The cages were 20 inches wide at the front, 12 inches deep, and 10 inches high. The top was hinged to permit introduction and removal of the male. The rear half of the top was surmounted by a cylindrical release can in which a receptive female was placed just before the test. The bottom of the can was hinged and could be dropped by raising a counterweight.

Three observation cages were fastened to a rack in such a position that they could be watched simultaneously by an observer seated approximately 6 feet away. At the observer's side were three counterweights controlling the trap-floors of the release cans.

The observer also operated three microswitches manually, each switch activating a separate marker which recorded on waxed paper pulled by a constant speed kymograph. One inch of paper passed under the markers every 12.6 seconds. The observer recorded different types of sexual responses made by the subjects by closing the microswitch for varying lengths of time. With this arrangement it was possible to observe and record three mating tests simultaneously. Later, when the kymograph records were analysed, the frequency and timing of response could be determined.

Testing Procedure

Incentive females were brought into heat by injections of estrogen and progesterone and their receptivity was pretested by noting their reactions to a non-experimental male. The dosages and timing of injections were those used in earlier studies (Beach & Holtz-Tucker, 1949). Only fully receptive females were used in the experimental tests. No attempt was made to pair each male with the same female in different tests.

All testing was done during the dark phase of the 24-hour cycle. Before the start of a test a receptive female was put in the release can and the experimental male was placed in the observation cage and allowed 5 to 15 minutes for cage adaptation. The longer period was used in early tests, but as the experiment progressed shorter periods of adaptation proved sufficient. At the end of the adaptation period the kymograph motor was started and the incentive female was dropped into the cage by lifting the counterweight.

Records made during the test included the number of intromissions, incomplete attempts at intromission, and ejaculations. The time at which each event occurred was automatically indicated by virtue of the fact that the recording paper moved at a constant speed.

All tests were continued until the criterion of sexual exhaustion was met. A male was judged to be exhausted when he allowed 30 minutes to elapse without mounting the incentive female. Males that failed to mount at least once during the first 30 minutes of a test were removed from the cage and scored as negative for that day.

Scheduling of Tests

Measuring exhaustion and recovery. The original design was based upon Schwartz's finding that 6 days of rest yielded full recovery from sexual exhaustion in male rats. Accordingly, 3 tests were scheduled to occur 6 days after complete exhaustion. One of these tests was followed by a retest after a 24-hour interval. Another was followed by a test conducted after a rest of 3 days. Tests immediately preceding the 1-day and 3-day tests were in turn preceded by 6 days of rest.

Results obtained in the course of experimentation led us to believe that sexual recovery was not complete in our animals after a 6-day rest. Therefore when the randomized schedule was completed 11 rats were given a rest of 15 days and then retested. Since a few animals had not ejaculated in the test preceding the 15-day rest, the recovery period actually varied from 15 to 21 days in different individuals. Six rats were tested again 1 day later. If no copulatory behaviour occurred they were allowed another 2 days of rest and tested once more a total of 3 days after the 15-day test. The remaining animals were tested 3 days after the 15-day test.

Measuring level of performance at shorter intervals. Beginning from 9 to 16 days after the final exhaustion test 5 males were tested for sexual activity once a day for 13 days. These tests lasted 15 minutes from the first mount or 15 minutes from the introduction of the female if no mounting occurred. Most males of our strain will achieve at least one ejaculation within 15 minutes from the start of copulation if they copulate at all. Six other males plus 3 additional experienced copulators were given 11 or 12 similar tests at 48-hour intervals, i.e. every other day.

RESULTS

The Course of Exhaustion in Fully Rested Males

Because of evidence to be presented below we concluded that males which had not engaged in sexual activity for 15 days could be regarded as fully rested and capable of maximal copulatory performance. Therefore behaviour in the 15-or-more-day-tests was analysed to study the course of sexual exhaustion.

Eleven males were tested after a rest of 15-21 days. In this test 10 animals ejaculated 5 or more times. One rat failed to copulate in the first 30 minutes of the 15-day test and was also negative at 16 days. On the eighteenth day he ejaculated 7 times. The reasons for such failures are obscure, but it is not uncommon for a vigorous male to skip a test occasionally.

Gross measures of sexual capacity. In fully rested males the time needed to reach sexual satiation varied from 61.5 to 141.5 minutes. The mean duration of the test from the introduction of the receptive female to the occurrence of the male's final ejaculation was 89.2 minutes (Median = 82.8). Subtracting from this total the time elapsing before the male first began to mate (*mount latency*), and also the periods of sexual inactivity following all but the terminal ejaculation (*refractory periods*), we find that the average number of minutes spent in actual copulatory activity was 36.7 (Median = 23.1). The range was from 17.4 to 89.7 minutes.

In terms of ejaculatory performance the criterion of sexual exhaustion was reached after a mean number of 6.9 ejaculations (Median = 7.0). Individual variation was considerable and the total number of ejaculations ranged from 5 to 10 for different males. In order to achieve this number of ejaculations the subjects effected intromission an average of 41.7 times (Median = 39.0) and the range was 24 to 67.

In the case of 8 of the 11 males the final ejaculation terminated all sexual activity. Only 5 per cent. of all intromissions and 18 per cent. of all incomplete mounts occurred after the terminal ejaculation. These scores were contributed by 3 rats which resumed copulating but did not ejaculate again and finally met the criterion by failing to mount the female for a period of 30 minutes. These results gave us confidence in our selection of the criterion for sexual exhaustion. Apparently if a previously active male ceases copulating for 30 minutes he would be unlikely to attain another

ejaculation in that test even if left with the female for a longer period of time. One or more ejaculations occurred in 57 of all the exhaustion tests, and in 53 (90 per cent.) of these tests all mounting activity was terminated by the final ejaculation.

The average delay between the introduction of the receptive female and the male's first mounting response (*mount latency*) was 74.0 seconds (Median = 8.0). The mean was heavily influenced by negative tests for which an arbitrary mount latency of 1,800 seconds (30 minutes) was assigned. Since there were only 2 negative tests this practice did not affect the median. When negative tests were excluded, the range in individual mount latencies was found to be 2 to 580 seconds.

Mount latency applies to the first mount whether it be a completed copulation with intromission or an abortive copulatory attempt. *Intromission latency* refers to the time between introduction of the incentive female and the occurrence of the first complete copulation. In fully rested males this measure had an average value of 91.3 seconds when negative tests were included (Median = 12.0). Counting only positive tests the range of intromission latencies was 2 to 750 seconds.

Behavioural changes as exhaustion approaches. To analyse the nature of sexual exhaustion we studied the changes in behaviour preceding each successive ejaculation up to the point at which the criterion was met. As noted above, only the records of males rested for 15 or more days were used.

Some of the results are summarized in Table I, which represents the situation up to and including the sixth ejaculation. So few animals ejaculated more than 6 times that group scores are not representative past this point. The average number of intromissions preceding the first ejaculation was 10.64 and for succeeding ejaculations this value decreased with the result that the sixth ejaculation was preceded by an average of only 4.10 intromissions.

TABLE I
CHANGES IN PERFORMANCE ASSOCIATED WITH SUCCESSIVE EJACULATIONS

Serial number of ejaculation	Intromissions to produce ejaculation	Seconds to achieve ejaculation	Seconds to recover and resume mating
1		450	324
2	10.64	216	395
3	6.00	198	468
4	5.73	150	495
5	5.09	210	597
6	5.09	132	818
	4.10		

Time scores shown in Table I are based upon group medians. *Ejaculation latency*, or the time from the first mount to the occurrence of the ejaculation was 450 seconds in the case of the first, and 216 seconds in the case of the second ejaculation. There was some tendency for further decrease with successive ejaculations. There was, at the same time, a clear-cut increase in the refractory periods that followed successive ejaculations. Lengthening of the refractory period is a premonitory sign of sexual exhaustion. For example, in the one male which ejaculated 10 times, successive refractory periods were as follows: 330, 389, 432, 432, 244, 364, 465, 881 and 1,082 seconds. The tenth ejaculation was, of course, followed by an unmeasured refractory period which exceeded 30 minutes.

The Course of Recovery from Sexual Exhaustion

It will be recalled that males were tested at 1, 3, 6 and 15 days after sexual exhaustion. Sometimes a male failed to copulate and was retested at some interval not included in this series. Our analysis of the following behavioural data has been confined to results obtained on 70 tests conducted at the stated intervals. Seventeen other tests which occurred at 7, 8, 9, 10, 12, 18 or 21 days have not been used in the statistical treatment although results from them are represented in some of the figures.

Changes in the test as a whole. The percentage of tests in which each type of sexual response occurred at least once after various recovery periods is shown in Table II. Relatively little recovery appears to have taken place during the first 24 hours after exhaustion and the curve is still rising after 3 days of rest. From these data one cannot determine whether full recovery has been achieved at the end of 6 days because the number of 15-day tests is too small to yield a stable score. Other findings to be presented below show that 6 days of rest does not produce complete sexual recovery.

TABLE II
PERCENTAGE OF TESTS IN WHICH EACH TYPE OF RESPONSE OCCURRED AT
DIFFERENT POINTS ON THE CURVE OF SEXUAL RECOVERY

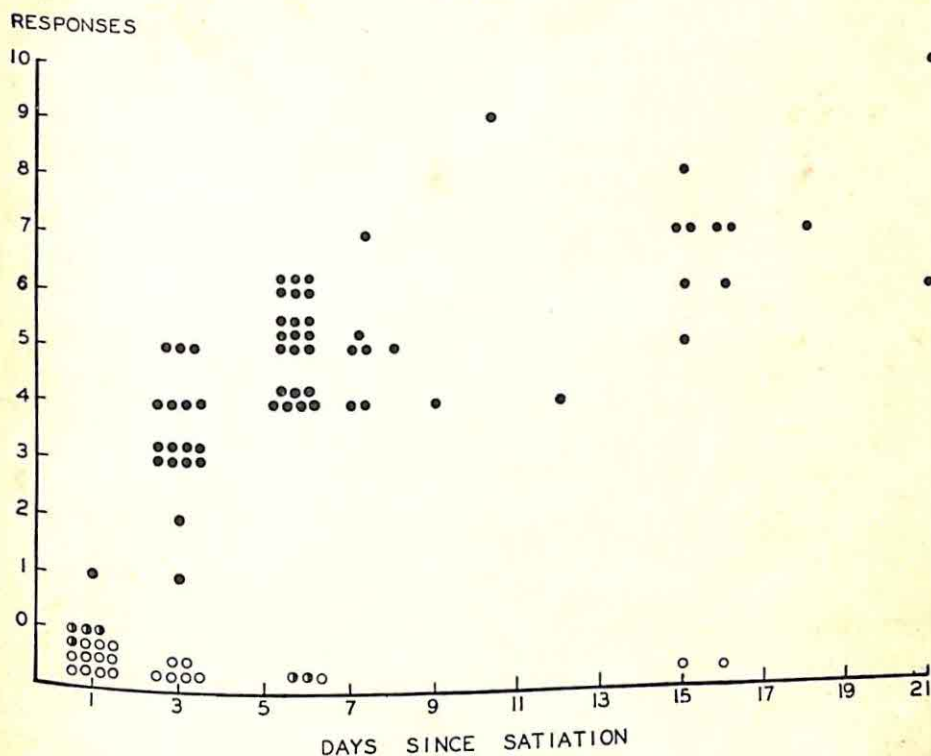
	<i>Days since sexual exhaustion</i>			
	1	3	6	15
Total number of tests	16	23	25	6
Per cent. with no mating activity	69	26	4	17
Per cent. with attempted copulation only ..	25	0	8	0
Per cent. with ejaculation	6	74	88	83

Ejaculation was achieved by only one male in the test conducted 24 hours after exhaustion, and this animal ejaculated but once. For those rats that reached ejaculation in 3-day tests, the median number of ejaculations to exhaustion was 3.0 ($R = 1-5$). At 6 days this value was 5.0 ($R = 4-6$), and after a 15-day rest the median number of ejaculations preceding exhaustion was 6.0 ($R = 5-8$). Figure 1 illustrates the number of ejaculations occurring in all tests including those that took place at intervals other than the 4 standard ones. A curve drawn through the median values for days 1, 3, 6 and 15 would show that sexual recovery is a negatively accelerated function and is probably complete at some point between the tenth and fifteenth days of rest.

TABLE III
MEAN NUMBER OF COMPLETE AND INCOMPLETE COPULATORY RESPONSES
NECESSARY TO PRODUCE EJACULATION

<i>Days since sexual exhaustion</i>	<i>Ejaculations</i>	<i>Copulations per ejaculation</i>	<i>Incomplete copulations per ejaculation</i>	<i>Combined</i>
3	58	8.19	9.10	17.29
6	109	6.20	7.14	13.34
15	33	6.27	5.97	12.24

FIGURE 1



Number of mounting responses or ejaculations to produce exhaustion in tests occurring after varying periods of rest. (Hollow circles represent no copulatory activity. Half-filled circles represent copulation but no ejaculation. Solid dots show the number of ejaculations occurring before the criterion of exhaustion was reached)

In general it appeared more difficult for partially rested males to achieve ejaculation than for other males which were more nearly recovered from sexual exhaustion. This difference is brought out in Table III. Animals that had rested only 3 days had to execute more copulatory responses to attain ejaculation than did males with 6 or 15 days of rest. Differences between the 6-day and 15-day tests were not pronounced, a finding which corroborates the conclusion that the curve of sexual recovery is negatively accelerated and that recovery is almost though not entirely complete at 6 days.

TABLE IV
AMOUNT OF ACTIVITY AND LENGTH OF TIME NECESSARY TO PRODUCE SEXUAL EXHAUSTION AFTER DIFFERENT PERIODS OF SEXUAL REST

From start of test to final ejaculation	Days since preceding exhaustion		
	3	6	15
Median Number of Minutes	58	60	87
Median Number of Intromissions	22	29	39
Median Number of Incomplete Copulations	22	17	30

Since the number of ejaculations tended to vary with the number of days of recovery it is to be expected that the total amount of sexual activity necessary to produce exhaustion would vary in like manner. This proved to be the case as is shown in Table IV. The number of intromissions and for the most part, of incomplete copulations preceding exhaustion increased in males tested after 3, 6 and 15 days of sexual inactivity. The length of time during which males remained sexually reactive increased similarly.

The median *intromission latency* (i.e. delay before the first completed copulation) in 1-day tests was 3,600 seconds. Comparable scores were 623 seconds for 3-day tests, 135 seconds for 6-day tests and 18 seconds for tests preceded by 15 days of sexual rest. This measure appears to be quite sensitive to sexual exhaustion and recovery. If we compare median *mount latencies* in tests conducted after different periods of rest the results are as shown in Table V. It will be recalled that mount latency refers to the first copulatory response regardless of whether intromission is achieved. Mount latencies like intromission latencies are clearly affected by the degree of recovery from a session of unlimited sexual activity.

TABLE V

EFFECTS OF VARYING PERIODS OF SEXUAL INACTIVITY UPON DELAY PRECEDING THE INITIATION OF SEXUAL CONTACT

	Days since sexual exhaustion				Grand median
	1	3	6	15	
Median Mount Latency in Seconds ..	1,800	429	30	33	116.5
Per cent. of Tests Above Grand Median	81	65	32	33	—
Per cent. of Tests Below Grand Median	19	35	68	67	—

Changes occurring as the test progressed. The course of sexual exhaustion after different periods of rest is reflected in Figure 2, which shows the time from the beginning of the test at which successive ejaculations occurred. The shape of the curves obtained after 3, 6 or 15 days of rest is similar. The positive acceleration reveals that successive ejaculations were spaced farther and farther apart as the exhaustion point was approached. The overall rate of decline is seen to be a function of the degree of recovery at the start of the test.

The function represented in Figure 2 is actually a combination of two measures: plug latencies and refractory periods. These are plotted separately in Figures 3 and 4. Here it is shown that refractory periods following successive ejaculations increased progressively for all groups and the rate of increase was slowest in males rested for 15 days and most rapid in rats tested after only 3 days of rest.

The curves for ejaculation latencies tell a different story. Time to achieve the first ejaculation was inversely related to duration of the rest period preceding the test. However, for all groups the remaining ejaculations were achieved in shorter times and no intragroup differences are evident. This was due to two factors. The number of complete and incomplete copulations preceding the first ejaculation was higher than that preceding subsequent ejaculations for all 3 groups, and the speed with which one mounting response followed another was increased after the initial ejaculation had been achieved.

FIGURE 2

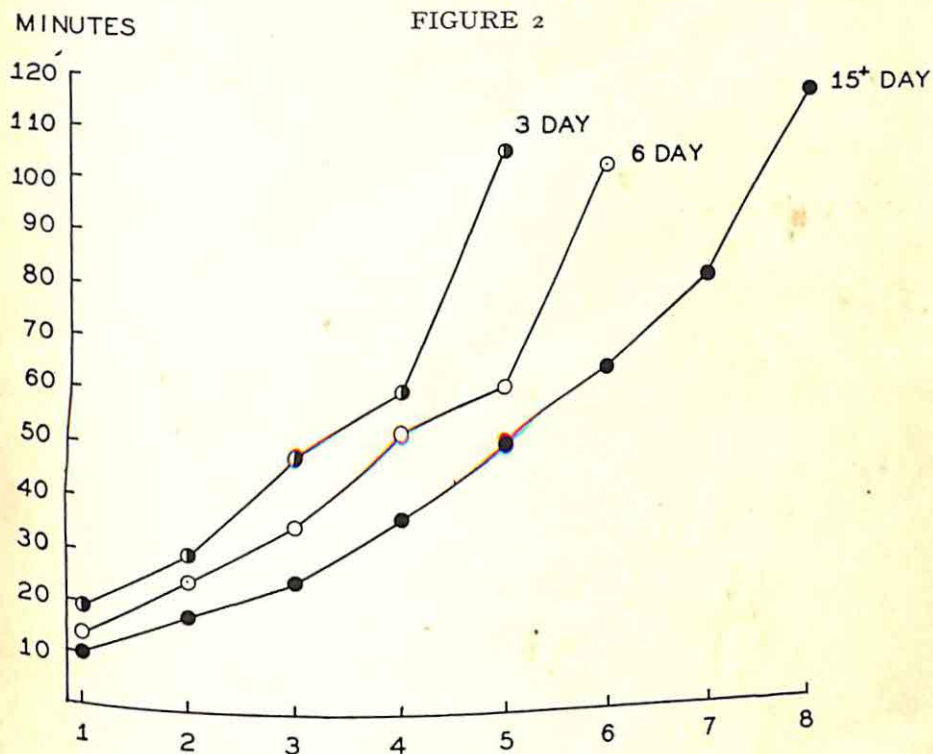
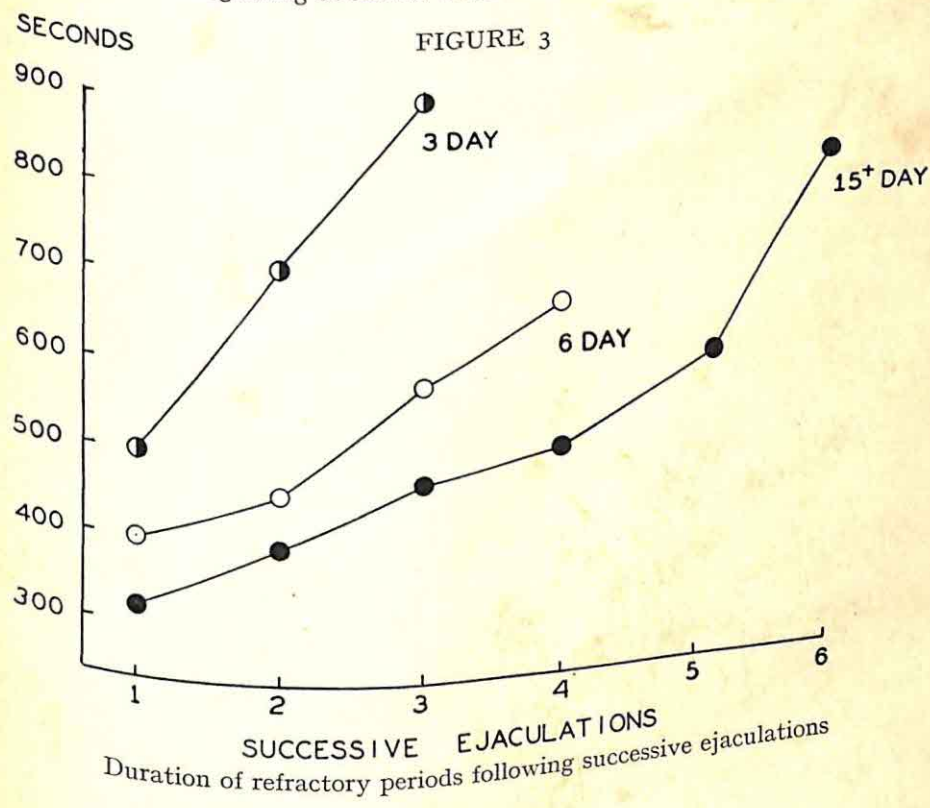
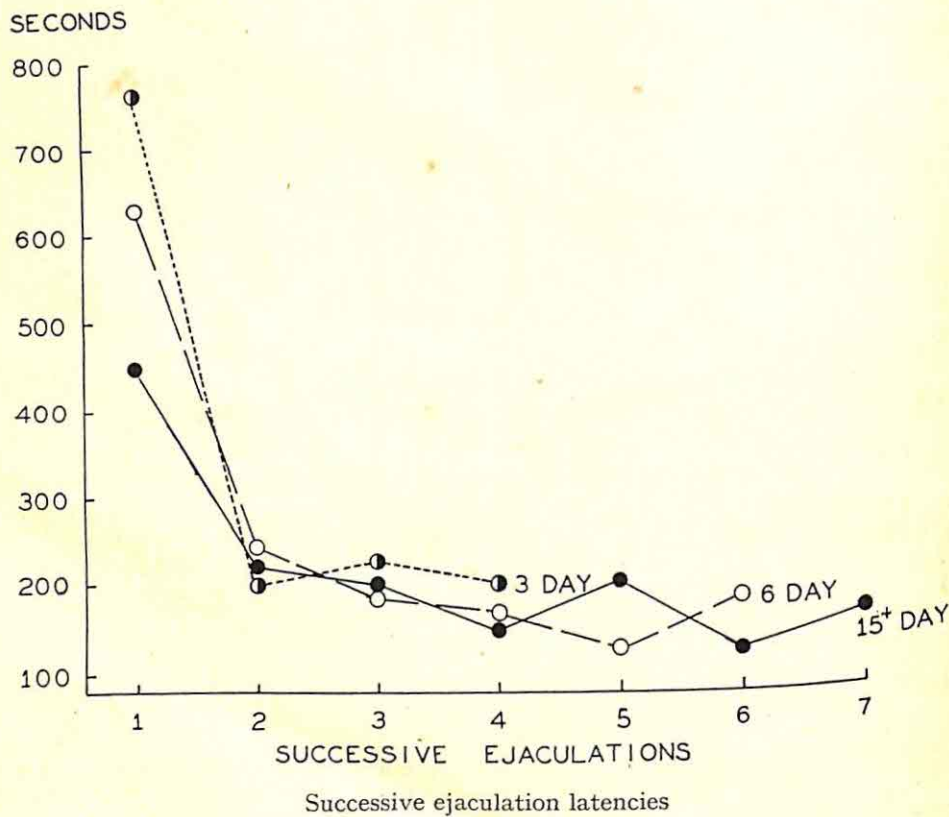


FIGURE 3



Short interval tests. It will be recalled that 9 rats were given 15-minute tests at intervals of every other day, and 5 males were tested once every 24 hours. Ejaculation occurred in 78 per cent. of the 107 tests conducted at 48-hour intervals and in 77 per cent. of the 65 tests given once a day. Ejaculation was present in only 83 per cent. of the tests made after a 15-day rest.

FIGURE 4



Median mount latency was 10 seconds on the 48-hour tests and 14 seconds on the 24-hour tests. In the 15-day tests the mount latency was 8 seconds. Intromission latency had a median value of 34 seconds for the 1-day, 17 seconds for the 2-day, and 12 seconds for the 15-day tests. The average number of intromissions to achieve one ejaculation was 10.4 in tests given daily, 11.4 for tests occurring on alternate days, and 10.6 in 15-day tests. Median ejaculation latency was 343 seconds in daily tests, 410 seconds in tests given every other day, and 450 seconds in tests preceded by a 15-day rest. Continuous graphs of sexual performance on successive short-interval tests reveal no consistent trends of behavioural change. The records of rats tested every day are somewhat more variable than those of males tested on alternate days, but there are no indications of progressive or cumulative sexual fatigue.

Considered in combination these findings present a reasonably clear picture. Male rats allowed to ejaculate once a day maintain sufficient sexual responsiveness to continue at a constant level of mating activity for considerable periods of time. Nevertheless the motivational level is probably below the maximum. Animals tested every other day have slightly greater sexual capacity but are not as responsive as fully rested males.

DISCUSSION

In an earlier publication the present writer postulated the existence of two physiological mechanisms involved in masculine sexual behaviour (Beach, 1955). The arousal mechanism (*AM*) mediates the awakening and increase of sexual excitement which eventuates in copulation. When the copulatory threshold is reached mounting and intromission occur and these responses are mediated by the copulatory-ejaculatory mechanism (*CM*). Coition produces a qualitative change in the sensory input resulting from genital stimulation. Ejaculation is the end product of a series of appropriately spaced intromissions.

Following an ejaculation the *AM* becomes temporarily unresponsive. As the *AM* regains its reactivity the male begins again to react to sexual stimulation; arousal occurs and the *CM* is once more thrown into action.

Effects of Sexual Exhaustion and Rest upon the AM

Sexual exhaustion. The two indicators of excitability of the *AM* are the intromission latency and the post-ejaculatory refractory periods. Changes in the latter measure as shown in Table I reflect the gradual change occurring in the *AM* during a period of unlimited coital behaviour. The duration of successive refractory periods is progressively increased. This is interpreted to mean that a series of ejaculations has a cumulatively inhibiting effect upon the responsiveness of the *AM*.

The length of time needed for recovery of excitability was increased by each successive ejaculation until finally this interval exceeded 30 minutes and the test was terminated. It is the ejaculation rather than total intromissions which depresses the *AM* because this effect is independent of the number of intromissions needed to produce orgasm.

Sexual rest. Failure to copulate on any given occasion may be due to a variety of reasons and excitability of the *AM* is certainly influenced by factors other than preceding ejaculations. Nevertheless the high proportion of negative tests occurring 24 hours after sexual exhaustion and the still appreciable number after a 3-day rest undoubtedly reflect in part submaximal recovery of the *AM* in some individuals.

More clear-cut evidence is provided by the median mount latencies recorded in positive tests occurring after 1, 3, 6 or 15 days of sexual rest. The respective scores in seconds were 1,800, 429, 30 and 33. An even more sensitive index to the state of the *AM* is the intromission latency. Median values for this measure were 3,600, 623, 135 and 18 seconds. It seems clear that recovery of excitability in the *AM* following sexual exhaustion is a negatively accelerated function and that a maximal state of responsiveness is probably not attained in less than one week.

Measurements of the rapidity with which males recover from the effects of successive ejaculations also reflect the effects of sexual rest upon the *AM*. The median time to resume copulating after the first ejaculation was 504 seconds in tests occurring after a 3-day rest, 397 seconds in 6-day tests and 324 seconds in tests following a rest of 15 or more days. The rate at which the duration of subsequent refractory periods increased was greatest in 3-day tests and slowest in 15-or-more-day tests as can be seen in Figure 3.

Excitability of the *AM* appears to be held slightly below maximum if males are allowed a single ejaculation every 48 hours, and the degree of depression is a bit greater if ejaculations are spaced 24 hours apart. Males that ejaculated once every 24 hours had a median intromission latency of 34 seconds. For those that experienced one ejaculation every other day this figure was 17 seconds, and males with a 15-day rest achieved intromission in 12 seconds.

Effects of Sexual Exhaustion and Rest upon the CM

Sexual exhaustion. Changes in the functional state of the *CM* are reflected in the amount of genital stimulation necessary to produce ejaculation and in the ejaculation latency. Present findings suggest that the first factor is reduced as a result of successive ejaculations. This "sensitizing" effect on the *CM* is most noticeable following the initial ejaculation. As shown in Table I, more than 10 intromissions preceded the first ejaculation in fully rested males, whereas the second orgasm occurred after only 6 such responses. Further reduction in intromissions necessary to produce succeeding ejaculations is indicated by the tabular data.

As measured in this fashion the excitability of the *CM* seems to remain high until the male ceases to make any sexual responses due to extreme loss of reactivity in the *AM*. One male ejaculated 10 times before reaching the criterion of sexual exhaustion. Each of the last two ejaculations was produced by 2 intromissions.

Table I also reveals that the length of time needed to reach ejaculation tends to decrease as the test progresses. This is, of course, due in part to the reduction in intromission frequency, but other factors to be discussed below may be involved.

Sexual rest. The amount of genital stimulation needed to produce the first ejaculation in a test varies according to the male's sexual condition. The total complete and incomplete copulations preceding the initial ejaculation was 23.2 in 3-day tests, 19.1 in 6-day tests and 17.4 in tests following a rest of 15 days.

The time needed to ejaculate for the first time varied inversely with the length of sexual rest. Median first plug latencies were 450 seconds in 15-day tests, 631 in 6-day tests and 756 seconds in tests conducted after 3 days of rest.

It is apparent that the excitability of the *CM* is low after an exhaustion test and recovers gradually as a result of sexual inactivity. At the same time there is a clear-cut tendency for successive ejaculation latencies to decrease in all tests regardless of time since the preceding exhaustion.

One ejaculation every 24 or 48 hours seems to have no detectable effect upon the *CM*. At any rate the number of intromissions necessary to produce one ejaculation was 10.4 in 24-hour tests and 11.4 in 48-hour tests as compared with 10.5 in 15-or-more-day tests.

A measure which as yet we have been unable to interpret in any meaningful fashion might be called *copulatory speed*. It is derived by dividing the number of intromissions to ejaculate into the ejaculatory latency. In effect this indicates the rapidity with which one copulation follows another. In the course of an exhaustion test the curve of copulatory speed first increases and then decreases. In tests of fully rested males the speed was 51.5 seconds to the first ejaculation, 34.9 seconds to the third and 52.7 seconds to the seventh ejaculation. In achieving their final ejaculation males tend, as noted above, to execute relatively few copulations but these are widely spaced in time. It is because of progressive changes in copulatory speeds that the copulatory latencies listed in Table I were said to reflect something more than a mere reduction in intromission frequency.

Estradiol benzoate (Progynon B) and progesterone (Proluton) used in this study were generously supplied by Dr. Edward Henderson, of Schering Corporation, Bloomfield, New Jersey.

REFERENCES

- BEACH, F. A. (1945). Hormonal induction of mating responses in a rat with congenital absence of gonadal tissue. *Anat. Rec.*, **92**, 289-292.
- BEACH, F. A. (1947). A review of physiological and psychological studies of sexual behaviour in mammals. *Physiol. Rev.*, **27**, 240-307.

- BEACH, F. A., and HOLZ-TUCKER, A. M. (1949). Effects of different concentrations of androgen upon sexual behaviour in castrated male rats. *J. comp. physiol. Psychol.*, **42**, 433-453.
- BEACH, F. A. (1955). Neural and Chemical Regulation of Behaviour. Wisconsin Symposium on Interdisciplinary research. In press.
- MILLER, N. E., and DOLLARD, J. (1941). *Social Learning and Imitation*. New Haven.
- SCHWARTZ, M. (1955). Instrumental and consummatory measures of sexual capacity in the male rat. *J. comp. physiol. Psychol.* In press.
- WARNER, L. H. (1927). A study of sex behaviour in the white rat by means of the obstruction method. *Comp. psychol. Monog.*, **4** (22).

DISTORTIONS IN IMMEDIATE MEMORY

BY

J. BROWN

From the Department of Psychology, Birkbeck College, University of London

A series of five pairs of abstract geometrical figures were presented to each subject. Unknown to the subjects, the members of each pair were identical. The second member was presented 10 seconds after the first member and the subject then judged whether a specified feature of the figure had changed in one of two possible directions. One group of 32 subjects were not told which feature of the figure would be critical until presentation of the second member: a further group of 32 subjects were given this information before presentation of the first member. Subjects in both groups were significantly consistent with one another in some of their judgements, thus indicating that distortions in remembering had occurred. It is argued that memory distortions of this type (for which previous evidence has been unsatisfactory) are of particular interest since, unlike other memory changes, they are difficult to ascribe to the constructive or inferential character of recognition and recall.

I

INTRODUCTION

Memory distortions may be held to arise either during retention and/or at the time of recognition or recall. For convenience, distortion arising during retention will be called retention distortion and distortion arising at the time of recognition or recall will be called recognition and recall distortion respectively. Both kinds of distortion have been alleged to occur. Retention distortion implies that some change takes place in the neural substrate of retention, other than simple decay. Koffka (1935) claimed that resolution of internal stresses in the memory trace, interaction between traces and communication with new perceptions all lead to such change in the trace during retention. Similarly, the schemata of Bartlett's theory (Bartlett, 1932) are held to undergo modification during retention; on this theory, however, the modification is ascribed to the effects of new perceptions only. Bartlett also emphasized that distortions arise from the reconstructive character of recall. Many others, especially psychologists in the Behaviourist tradition, have contended that all memory distortions are due to recall or recognition distortion.

It is certainly plausible to attribute many memory distortions to recall or recognition distortion. If the original material has been partially forgotten, explicitly or implicitly recall or recognition will involve a certain amount of inference. Such inference must often lead to distortions. Now the sources of information on which it can be based are: (i) what the subject still remembers of the material; (ii) what he remembers of his cognitive and affective reactions to the material (this source of information is not sharply distinguished from the first: the subject's reactions to the material are, in a general sense, his ways of coding it); (iii) his knowledge and beliefs as to what is probable. Hence the distortions which may be expected to result from the inferential element in recognition and recall include exaggeration of the relative importance of aspects of the material which are still remembered, changes based on inferences from affective and cognitive reactions and, finally, distortions in the direction of the conventional and expected.

Before distortions which are difficult to attribute to recall and recognition are considered, an argument for the occurrence of retention distortion which is peculiar to Bartlett's theory of remembering should be mentioned. The argument proceeds

from the character of recall distortions and might be expressed as follows. Recall distortions are typically in the direction of the conventional and expected, and thus reveal a knowledge of probabilities. Now knowledge of probabilities does not seem to depend primarily on remembering and reviewing relevant experiences, but seems to develop continuously from the succession of individual experiences, the majority of which are soon forgotten. This suggests that retention is mediated by the modifications which individual experiences produce in plastic schemata and not by individual memory traces. Since later experiences will further modify the schemata, retention distortion will occur. This conclusion only holds, however, if both the learning of probabilities and the retention of individual events are ascribed to schemata; it does not hold if the retention of individual events is ascribed to individual traces. And it is very doubtful whether it is possible to explain retention of individual events in terms of schemata alone (Oldfield & Zangwill, 1943; Gomulicki, 1953).

The main class of distortions which are not easy to attribute to recall or recognition distortion are those which have been reported in the remembering of abstract visual figures and which are claimed to be fairly constant from subject to subject. Thus if there is consistent partial closure of the gap in a circle, there is no obvious way in which this could be the product of inferential recognition or recall. A smaller gap is no more probable than a larger gap, for example. The interpretation of such changes will be further considered in the discussion. A prior question is whether such changes do in fact occur. For one reason or another, the evidence has been unsatisfactory. Most of the investigators who reported positive results used the method of reproduction. This method has serious defects: classification of the reproductions is problematical and the reproductions are influenced by drawing habits (Hanawalt, 1937; Woodworth, 1938). Hebb & Foord (1945) and George (1952) used a recognition method and obtained little evidence of consistent memory changes. However, in their recognition test, the subject had to select what he believed to be the correct figure from a series of similar figures: any changes may have been masked by the influence of central tendency of judgement and by the influence of the further memory traces established by the first figures examined. The most satisfactory evidence of Gestalt-type changes was obtained by Irwin & Seidenfeld (1937), using the method of comparison; although Hanawalt (1952) in a similar experiment, failed to repeat most of their results. Their method was to present a series of stimulus figures one at a time to a group of subjects. Three minutes later, they presented a second series which, unknown to the subjects, was identical to the first series. As each figure of this second series was presented, the subjects were asked specific questions as to how it differed from the similar figure of the first series. This method avoids both the faults of the reproduction method and is essentially the one adopted in the present experiment. In the form in which they used it, however, the method is open to the following criticisms. Firstly, if a subject's attention is directed to some particular feature of a figure, this may alter the appearance of the figure—very strikingly, if it is an ambiguous figure or a puzzle picture, for example. Hence a perceptual effect of this kind might explain their results, since the attention of their subjects was not directed to the features about which judgements were required until the second presentation of the figures. Secondly, since the series of figures was presented as a whole, errors of judgement about individual figures could arise through intra-serial confusion.

II

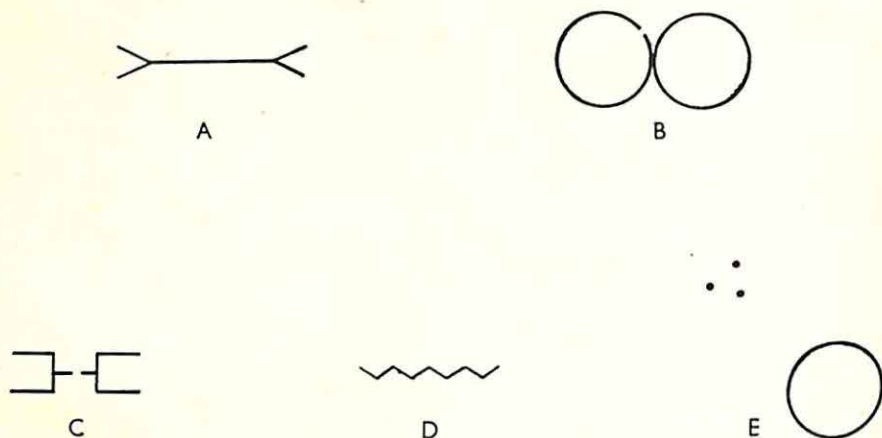
METHOD

The method of comparison was used under two separate conditions. Under one condition, the subject was not told until the presentation of the second member of each

pair of stimulus figures which feature of the figures was to be compared. Under the second condition, this information was given before the presentation of the first member of the pair. This condition is a check on the possible perceptual effect mentioned as a criticism of Irwin and Seidenfeld's experiment. The second member of each pair was presented 10 seconds after the first member.

A series of five pairs of stimulus cards was prepared. The cards of each pair were identical, apart from minor blemishes. Each pair carried a different stimulus figure. The stimulus figures (reduced in size) are shown in Figure 1. They were drawn in black ink on the cards which were of white board, 20 cm. square. The size of each stimulus figure was such that it was well away from the edges of the card: for example, A was 10 cm. long. Each subject was tested separately and was seated about 4 feet from the experimenter.

FIGURE 1



The Stimulus Figures.

Condition I. The subject was told, "I am going to show you a card for three seconds. It has on it a simple geometrical design such as a triangle and a line. After an interval of ten seconds, I will show you a second card bearing a similar design to the first one. I will ask you some question about how the second card differs from the first. As you will be seeing the cards in succession, and not side by side, the difference may not be very obvious to you—but if you have an impression of difference, I want you to tell me which way it goes." The terminal statement was intended to encourage the subject to report any difference he thought he saw and to prevent him from suspecting that the cards were in fact the same. The procedure outlined to the subject was then carried out using the first pair of identical cards. Immediately before being shown the second card, the subject was asked one of the following questions. With A, "Is the centre line longer or shorter?" with B, "Is the gap in the circle larger or smaller?"; with C, "Is the break in the centre line larger or smaller?"; with D, "Are the angles of the figure flatter or sharper?"; with E, "Are the dots closer to one another or further apart?" The second card was presented until the subject made his judgement: this was usually within 2 or 3 seconds. After his judgement had been recorded, the remaining pairs of cards were presented using the same procedure.

Half the subjects were shown one card of each pair first and half were shown the other card first, in order to balance out any biases introduced by slight differences in the cards. The order of presentation of different pairs of cards was varied from subject to subject but not systematically.

Condition II. This condition differed from the first in that the subject was told what question would be asked about the second card immediately before presentation of the first card.

One group of 32 subjects was tested under condition I and a second group of 32 subjects under condition II. The subjects were male and female undergraduates and research workers.

III RESULTS

The question is whether there is a significant tendency with some of the stimulus figures to report a change in one rather than the other of the two possible directions. In Table I, f denotes the number of subjects who reported a change in the more popular direction; n denotes the number who reported a change in either direction.

TABLE I
CHANGES REPORTED

	Figure	Judgement	f	n	χ^2	Significance
Condition I	A	Centre line is shorter ..	21	29	4.97	$P < 0.05$
	B	Gap is smaller ..	20	27	5.33	$P < 0.025$
	C	Break is smaller ..	17	24	3.37	not significant
	D	Angles are flatter ..	28	32	16.53	$P < 0.01$
	E	Dots are closer ..	14	25	0.17	not significant
Condition II	A	Centre line is shorter ..	21	26	8.65	$P < 0.01$
	B	Gap is smaller ..	11	22	0.00	—
	C	Break is smaller ..	15	25	0.67	not significant
	D	Angles are flatter ..	18	26	3.12	not significant
	E	Dots are closer ..	6	24	5.04	$P < 0.025$

(χ^2 tests whether f/n differs significantly from $\frac{1}{2}$ —see text)

A χ^2 test (with Yates correction) shows whether the proportion f/n differs significantly from the expected value of $\frac{1}{2}$ on the null hypothesis. It will be seen that three out of the five values of f/n were significant under condition I and two out of five under condition II. The results were exceptionally consistent for figure D under condition I: all 32 subjects reported a change and 28 reported that the angles were flatter. The trend in the results was much less marked under condition II, when the subjects knew before the first presentation of D which feature would be critical: many of these subjects reported that they made estimates of the size of the angles in degrees. With figure A, the trend in the results was actually more marked under condition II. With figure E, a small non-significant trend was obtained under condition I and a significant trend in the opposite direction was obtained under condition II. Many subjects reported that they had not attended to the relevant feature (distance of the black dots from one another) of this figure under condition I when it was first presented: instead they had studied the distance of the black dots from the circle. None of the differences between the trends obtained under the two conditions are actually significant on a χ^2 test. More judgements of "no change" were given under condition II (37 out of 160) than under condition I (23 out of 160).

It may be asked whether subjects had any tendency to choose either the first or the second of the two alternatives stated in the question about each figure, since this would influence the results. Examination of Table I shows that the most frequent judgement corresponded to the first alternative with figures A and B and to the second judgement with figures C and D; while with E it varied with the condition. Thus no such tendency can be observed.

IV DISCUSSION

The subjects were led by the instructions to expect changes in the stimulus figures, but they were not led to expect any particular directions of change. Yet, with some

of the figures, there was a significant tendency for subjects to report a change in one rather than the other of the two possible directions. This implies either that these figures were perceived differently on their second presentation or that memory distortions had occurred. A possible source of perceptual change is a change in perceptual attitude between the two presentations. This is unlikely to explain the results, however, since significant tendencies were obtained even in the group of subjects who were told which feature would be critical before the first of the two presentations and for whom change in perceptual attitude should therefore have been minimal. The tendencies were somewhat less marked for this group. This was not unexpected since there is evidence that a selective learning set can have a considerable influence on the strength of the memory trace established at the time of perception (Brown, 1954)—and hence, indirectly, on the likelihood of memory distortion. The results bear some resemblance to figural after-effects and might be regarded as a species of perceptual change for this reason. However, the initial inspection period of each figure was only 3 seconds, so that it is doubtful whether the two phenomena are related. The conclusion therefore is that the results demonstrate memory change rather than perceptual change.

The similarity of the experiment to time-error experiments will have been noticed. It differs from such experiments partly in the complexity of the stimuli which the subject is asked to compare and partly in the absence of any obvious criterion for classifying a judgement as a positive or a negative error. The results in time-error experiments are apt to be unstable. Hanawalt (1952) maintains that this is also true of results on memory change obtained by the method of comparison. Only further work can show how far the present results are a function of the particular experimental situation employed.

The experiment seems to demonstrate the occurrence of distortions in the remembering of abstract visual material, by a method less open to criticism than methods used previously. The fundamental problem this raises is whether such distortions provide evidence for retention rather than recognition distortion. As argued in the Introduction, they are less easy to attribute than are most memory distortions to recognition distortion (or to recall distortion, if a recall test has been employed). But it does not follow that they *cannot* be attributed to recognition distortion. There are at least two ways in which such distortions might be produced by recognition distortion. Firstly, although there is no apparent reason why subjects should expect particular changes in any of the figures, subjects may have had unsuspected preconceptions on the matter. Secondly—and this perhaps is more likely—the distortions may have been produced in the following way. The same figure may tend to evoke similar coding reactions in different subjects. These reactions may then steer recall. For example, a verbal description such as "figure with long centre line" might tend to be given for figure A. If such a description, although appropriate for A, would be even more appropriate if the centre line of A were still longer, this would tend to make subjects judge the centre line of A as shorter on the second presentation. This theory could be tested by first asking subjects how they had classified a given figure and then studying to which of a series of similar figures, together with the original figure, the description was regarded as being most appropriate.

Even if it proves difficult to attribute these distortions to recognition distortion, this would constitute only a negative reason for attributing them to retention distortion. We may now consider possible reasons of a positive character for attributing them to retention distortion. Firstly, if it were shown that these distortions occur with all subjects, this might be explained more easily if they were attributed to retention distortion. For retention distortion could be presumed to depend on the

intrinsic properties of the neural trace, whereas recognition distortion is influenced by many factors peculiar to the individual. In the present experiment, only significant *trends* were demonstrated: some subjects reported no change with some figures and others reported a change in the direction opposite to the general trend. However, most subjects were uncertain in making their judgements. This implies that the changes were only small ones, and a small change is not likely to be effective *every* time a judgement is made, as experiments on thresholds show. The problem of the consistency of these changes therefore requires further study. Secondly, if the distortions were found to represent a further emphasis of unlearned principles of organization in perception (as Gestalt psychologists have maintained), this too, might seem to be more consonant with the view that they are retention distortions. Whether or not this is so, it may be remarked that the changes obtained in the present experiment cannot be predicted from the Gestalt laws of perception. There was no closure of the gap in the remembering of figure B, for example. With figure A, the centre line of this Muller-Lyer figure was remembered as even longer than it looked (since it was judged to be shorter on the second presentation): this provides an interesting example of the further emphasis of a perceptual distortion in remembering—but it is not a distortion which exemplifies the Gestalt laws in any obvious way, nor do we know whether it is an unlearned distortion. Thirdly, it has been held that if memory distortions can be shown to be progressive, this would tend to show that they originate in retention distortion. This contention is unconvincing, however, since progressive distortion can be the product of increasingly inferential recognition or recall, as Postman has recently shown (Postman, 1954).

Thus the memory distortions obtained in the present experiment, while they do not provide positive evidence for retention distortion, do at least constitute a challenge to the theory that all memory distortions are due to recall or recognition distortion. The method demonstrated in the experiment may prove useful in future work on the problem of memory change.

I wish to express my gratitude to Professor G. C. Drew, of Bristol University, in whose department I was working when this experiment was performed.

REFERENCES

- BARTLETT, F. C. (1932). *Remembering*. Cambridge.
- BROWN, J. (1954). The nature of set-to-learn and of intra-material interference in immediate memory. *Quart. J. exp. Psychol.*, **6**, 141-148.
- GEORGE, F. H. (1952). Errors of visual recognition. *J. exp. Psychol.*, **43**, 202-206.
- GOMULICKI, B. R. (1953). The development and present status of the trace theory of memory. *Brit. J. Psychol. Mon. Suppl.* No. 29.
- HANAWALT, N. G. (1937). Memory trace for figures in recall and recognition. *Arch. Psychol.* New York, **31**, No. 216 (quoted in Woodworth, 1938).
- HANAWALT, N. G. (1952). The method of comparison applied to the problem of memory change. *J. exp. Psychol.*, **43**, 37-42.
- HEBB, D. O. and FOORD, E. N. (1945). Errors of visual recognition and the nature of the trace. *J. exp. Psychol.*, **35**, 335-348.
- IRWIN, F. W. and SEIDENFELD, M. A. (1937). The application of the method of comparison to the problem of memory change. *J. exp. Psychol.* New York.
- KOFFKA, K. (1935). *Principles of Gestalt Psychology*. New York.
- OLDFIELD, R. C. and ZANGWILL, O. L. (1943). Head's concept of the schema and its application in contemporary British psychology. III. Bartlett's theory of memory. *Brit. J. Psychol.*, **33**, 113-129.
- POSTMAN, L. (1954). Learned principles of organization in memory. *Psychol. Mon.*, **68** (3), No. 374.
- WOODWORTH, R. S. (1938). *Experimental Psychology*. New York.

MISCELLANEA

A NOTE ON IMMEDIATE MEMORY

BY

O. L. ZANGWILL

From the Psychological Laboratory, University of Cambridge

In the course of an extensive inquiry into verbal memory processes (Gomulicki, 1952), some data were collected which appear to throw fresh light on the nature of immediate memory. In particular, it was found that many of the varieties of constructive change defined by Bartlett (1932) in connexion with long-term recall could be traced in the immediate reproduction of passages so brief as barely to exceed the memory-span. As changes of this nature do not appear to have been previously described in connexion with short-term recall some illustrative extracts from our material will be given. A full analysis along somewhat different lines has been presented elsewhere (Gomulicki, 1952).

In Gomulicki's experiments, narratives varying in length between 15 and 200 words were read aloud to 50 subjects, who were required to reproduce each passage orally as accurately as possible immediately after it had been read. The subjects (of whom half were women) were University graduates and undergraduates aged between 20 and 30. All subjects were tested individually and their reproductions recorded *verbatim*. In the present note, we shall be concerned only with the results given in the case of two passages, respectively 22 and 44 words in length. Despite their brevity, striking and consistent changes were observed in the reproductions which merit brief analysis.

MATERIAL AND RESULTS

Passage A. This passage read as follows: "A surprise awaited us. The children were up, in nightclothes and dressing gowns. Helped by Cook, they had prepared the breakfast-table."

The principal results were as follows:

- (1) Despite its brevity, this passage was reproduced *verbatim* by only 5 subjects.
- (2) Analysis of the reproductions showed that certain parts of the passage were reproduced substantially unaltered whereas other parts underwent significant change.
 - (a) "A surprise awaited us" was reproduced unaltered by 45 subjects, with minor changes by 4 and was omitted by 1 subject.
 - (b) "The children were up" was reproduced *verbatim* by all 50 subjects.
 - (c) "in nightclothes and dressing gowns" was reproduced unaltered by only 22 subjects. "Nightclothes" was omitted by 11 and "dressing gowns" by 4 subjects. Five subjects substituted "nightgowns" for "nightclothes" and 4 condensed "nightclothes and dressing gowns" into "nightgowns." Three subjects omitted the whole phrase and one substituted for it the words "completely dressed."
 - (d) "Helped by Cook they had prepared the breakfast-table" was reproduced with complete accuracy by only 9 subjects. Six altered the wording in such a way as to state that it was the children who had helped the cook and 4 attributed preparation of the breakfast-table exclusively to her. Examples are "The children were up, in their dressing gowns, to help the cook with the breakfast table"; "... and helped the cook to prepare the breakfast for us"; "... and the Cook prepared the breakfast-table." Two subjects stated that the cook had helped the children to get dressed. In brief, 12 subjects materially altered the sense of the passage by reversing the role of the cook and the children or by similar conventionalization.
 - (e) "breakfast-table"; 18 subjects whose reproductions were otherwise substantially correct omitted "table." Variations on the last phrase were "prepared the breakfast and laid the table" (3 subjects), "laid breakfast" (2 subjects), "cooked the breakfast" and "set the breakfast-table" (1 subject each).

The distribution of correct reproductions, omissions and changes in the different parts of this passage are shown in Figure 1.

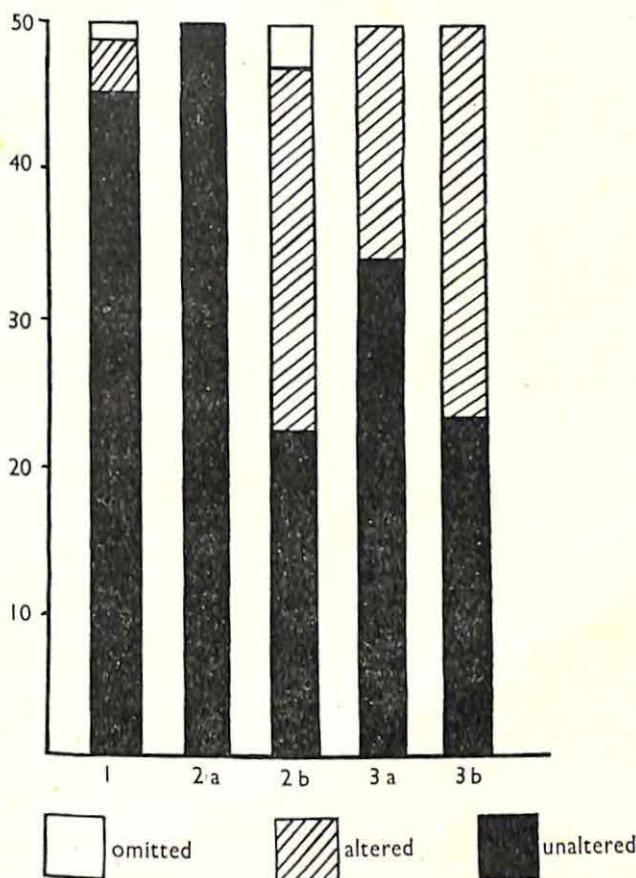
Passage B. This passage read as follows: "Mack, aged ten, is a fine carpenter. He did Uncle Bob a pipe for Easter and carved out a fine wooden doll, all in one piece, for his young nephew. He has since been working on a hutch for his ferrets to live in."

The principal results are as follows:

- (1) No subject reproduced this passage with word-perfect accuracy.
- (2) Certain inessential words and phrases were consistently omitted, e.g. the phrase "all in one piece" was dropped by 39 subjects.

FIGURE 1

Passage A



ANALYSIS OF REPRODUCTIONS OF PASSAGE A

1. "A surprise awaited us"
- 2a. "The children were up"
- 2b. "in nightclothes and dressing gowns"
- 3a. "Helped by Cook"
- 3b. "they had prepared the breakfast-table"

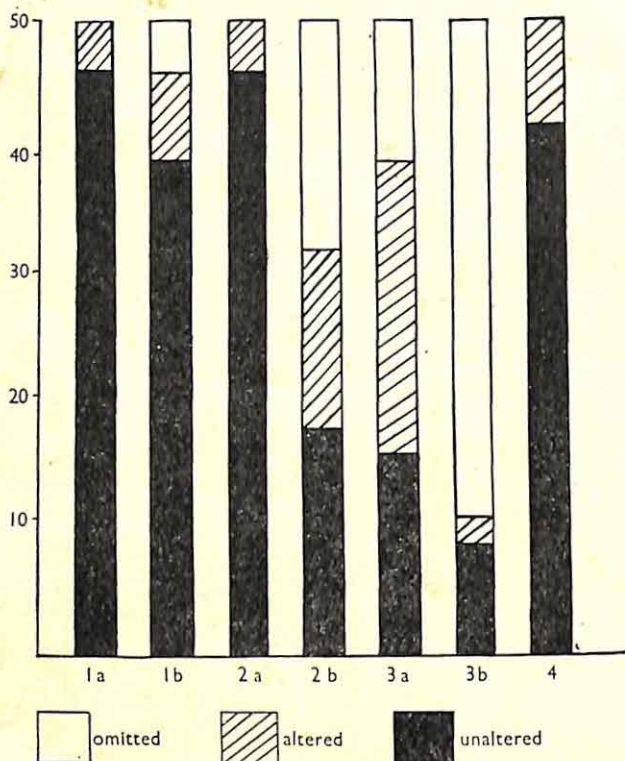
(3) Conventionalization was very marked. This affected several items in the passage which might be held to conflict with habitual expectations. First, presents are seldom given by children to adults at Easter; secondly, dolls are conventionally regarded as unsuitable presents for little boys; thirdly, boys as young as 10 do not often have nephews, and fourthly, ferrets are not often kept as pets. These unusual features were dealt with unwittingly in a variety of ways which repay brief analysis;

(a) *The Easter gift:* All reference to Easter was omitted by 18 subjects and 2 changed the phrase to "at Easter," thus minimizing the significance of the pipe as a gift. Substitutions comprised "for his birthday" (6 subjects) and "for Christmas" (5 subjects). One

subject stated simply that the pipe was given "as a present." Hence conventionalization is in this case achieved by omission in 18 cases and by various forms of substitution in 14, making 32 in all.

FIGURE 2

Passage B



ANALYSIS OF REPRODUCTIONS OF PASSAGE B

- 1a. "Mack is a fine carpenter"
- 1b. "aged ten"
- 2a. "He did Uncle Bob a pipe"
- 2b. "for Easter"
- 3a. "and carved out a fine wooden doll for his young nephew"
- 3b. "all in one piece"
4. "He has since been working on a hutch for his ferrets to live in"

(b) *The doll and the nephew*: The entire clause "and carved out a fine wooden doll, all in one piece, for his young nephew" was omitted completely by 11 subjects. Two retained the doll but omitted all reference to the nephew. Substitutions occurred in 14 cases, comprising "sister" (9 subjects), "cousin" (4 subjects) and "brother" (1 subject). One subject substituted "dog" for "doll" (possibly a mishearing). In 2 cases, "nephew's ferrets." Four subjects displayed a kind of *partial amnesia* for this part of the passage, e.g. "has since been engaged on something," "something about a doll in the middle," "something about a nephew." In short, this part of the passage was omitted in whole or in part by 17 subjects, and conventionalized, either by substitution or displacement, by 16. Only 16 subjects reproduced this clause essentially unaltered.

(c) *The ferrets*: "Rabbits" replaced "ferrets" in 7 reproductions and there were 3 omissions, in 2 cases with a sense of something having been forgotten (*partial amnesia*). The distribution of correct reproductions, changes and omissions of the different parts of the passage are shown in Figure 2.

DISCUSSION

These very limited results demonstrate that abstraction and constructive change may be detected in the immediate recall of very short narrative passages. Although some economy in reproduction was anticipated, the extent of conventionalization (Bartlett, 1932) in immediate recall came as a distinct surprise. This process (in short-term memory at least) would appear to operate in four principal ways: (1) by omission, as when a word or phrase whose significance is at variance with habitual expectation is simply dropped out, with or without awareness of something having been forgotten; (2) by minor alterations in word order or syntax, which may produce significant change in sense with minimal change in wording; (3) by substitution, as when an imported word or phrase replaces the original; and (4) by displacement, as where a word or phrase is transposed from one place to another within the passage. In all cases, conventionalization appears to occur wholly without the subject's knowledge or intention, i.e. as an unwitting process.

These preliminary observations might suggest that conventionalization has two closely related aspects, which may be termed the *factual* and the *verbal*. Factual conventionalization occurs when issues of fact or meaning are altered in such a way as to accord more closely with everyday experience. In the present material, it is well shown in the reversal of roles between the cook and the children in Passage A and the frequent substitution of "sister" for "nephew" in Passage B. Verbal conventionalization occurs when information is distorted as a result of its reformulation in accordance with current linguistic usage. Thus the omission of "table" in Passage A almost certainly derives from a preferred form of words ("prepare the breakfast" or "lay the table") and not from inconsistencies of a factual order. Although the conventionalizations of fact are the more dramatic, the possibilities of distortion due to formal linguistic usage certainly merit further exploration.

In conclusion, it may be stressed that the changes observed in these reproductions reflect active processes of reorganization in perception and recall. Omission and increased brevity of reproduction with increasing length of material can certainly not be attributed to passive decay; nor can conventionalization be treated as a special and merely incidental source of distortion. In our view, both abstraction and conventionalization are to be regarded as dynamic processes in precisely the same sense as is the process of repression postulated by Freud. Indeed it may be hazarded that Freudian "primary repression" is no more than a special case of what we have called *conventionalization by omission* and that it is subject to the same principles as govern memory-activity in general. At all events, the active character of immediate recall demands recognition by all concerned with the theory of memory, more especially, perhaps, by those whose approach to the subject is by way of communications theory (*cf.* Oldfield, 1954).

REFERENCES

- BARTLETT, F. C. (1932). *Remembering*. Cambridge University Press.
 GOMULICKI, B. R. (1952). "Recall as an abstractive process." D.Phil. Thesis, University of Oxford.
 GOMULICKI, B. R. (1956). "Recall as an abstractive process." *Acta Psychol.*, **12**, 77-94.
 OLDFIELD, R. C. (1954). "Memory mechanisms and the theory of schemata." *Brit. J. Psychol.*, **45**, 14-25.

I am greatly indebted to Dr. B. R. Gomulicki, who performed these experiments, for his kindness in permitting me to make use of his material.

CORRIGENDUM

In "Experimental Determination of Luneburg's Constants σ and K " by A. Zajackowska, Vol. VIII, Part 2, page 77, FIGURE 6: in case of Observer T.K., $A = 1.18$ (not 1.22 as misprinted); in case of Observer H.S., $a' = 0.86$ (not 3.86).

ANNOUNCEMENT

New Subscription Rates

for the

QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY

The subscription to the *Quarterly Journal of Experimental Psychology* has remained at 30 shillings per volume since it first appeared in 1948. Since then, however, costs have risen steadily, and a revision of rates is now necessary. The new rates, which will apply to Volume IX and thereafter, will be as follows:—

Ordinary Subscription per volume

of 4 numbers £2 10s. 0d.
(post free)

Special Subscription for under-
graduates and graduates of less
than three years standing per

volume of 4 numbers £1 5s. 0d.
(post free)

After December 31st, 1956, the cost of any single number will be 16/- (post free) and of any volume, £2 10s. 0d. (post free).

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY

Vol. VIII

1956

Part 4

SUCCESSIVE RESPONSES TO SIMULTANEOUS STIMULI

BY

D. E. BROADBENT

From the Applied Psychology Research Unit, Cambridge

Previous work had shown that, when a memory-span experiment is performed with half the items presented to one ear, and half simultaneously to the other, a certain order of response appears. Either one ear or the other is dealt with first, and then the remaining items produce a response afterwards. The present results extend this finding to the eye and the ear rather than the two ears, and also to two voices distinguished by their frequency characteristics. For the latter condition, it is also shown that alternation of attention can take place at a speed faster than between the two ears. The effect thus appears to be a general one, and not merely a peculiarity of binaural hearing.

I

INTRODUCTION

During a recent investigation on another topic (Broadbent, 1954) a finding was made which deserved further work. This finding was, that if two different series of stimuli were applied at the same time each to a separate one of the listener's ears, they could both elicit responses, but would do so one after the other. In the experiment actually reported, a set of three digits such as 713 were applied to one ear, and a different set such as 256 applied to the other ear. The subject's natural response was either 713256 or else 256713, but never 721536. Deliberate instructions to respond in the latter order produced a very low proportion of correct responses compared with the former orders. Since, then, all the numbers of one stream were responded to before any of the other stream, it follows that the purely mechanical difficulty of saying two digits at once is not relevant, as it might have been if only one digit were used on each ear. A delayed digit is not merely delayed until one other word has been spoken, but until all the digits on the other ear have been spoken. The delay is related to the sensory channel involved; it is perceptual rather than motor.

Two main lines of interest spring from this work. On the one hand, as has been pointed out by Mowbray (1953, 1954), the extent to which more than one process can go on simultaneously within the nervous system is an important one for such theories as that of Hebb (1949). The possibility that more than one complex stimulus can be perceived at once is supported by Hebb, and fits well into his general theory, although probably not rigorously implied by it. His theory holds that different stimuli activate physically different elements in the brain, and more than one might therefore be active simultaneously. A "mass action" view might hold rather that the whole cortex is involved in response to one stimulus, and so cannot simultaneously be controlled by another stimulus, though again the deduction is not likely to be

crucial. Mowbray (1953, 1954) provided evidence that, when the eye and ear are given complex stimuli at the same time, one or other may give rise to response, but not both. He had, of course, reduced the opportunities normally present in every-day life for rapid alternation between modalities, and his data supported the unitary nature of attention.

But it is probably a false antithesis to suggest that attention is purely unitary or completely divisible. In the effect we are considering, for example, it is clear that some mechanism must be involved which can only deal with different items successively. If this were not so, the order of the response could be more flexible. Yet it is also true that some other mechanism must be capable of receiving the various items simultaneously: otherwise half the items would be lost for ever. Such a compromise is perhaps to be expected from a nervous system which is known to have specialized parts and yet operates as an integrated whole. It is, therefore, of interest to find whether this preferred order of response occurs only with binaural presentation or whether it applies generally to simultaneous stimulation.

The ears might well differ from the eyes by having some temporary storage system which allows this effect to appear: and it might also be that the effect is connected with the division of the cortex into two hemispheres, in each of which one ear produces greater activity than in the other. Before drawing any general conclusion about the mechanism underlying successive response to simultaneous stimuli, it is necessary to vary the sensory channels employed: and this is the purpose of the present paper.

II

EXPERIMENT I

EYE AND EAR

The first variation studied was the use of eye and ear as the two channels of simultaneous stimulation. Mowbray, as has been said, found that information delivered to one of these senses was neglected if information from the other was simultaneously being used. The experiment was intended to find, firstly, whether Mowbray's conclusion still holds with rather different material. Secondly, if response to simultaneous stimuli proved possible, was it carried out by delaying all the information on one of the sensory channels?

(a) *Method.*

The same task was used as that employed by Broadbent (1954): namely subjects were presented in various ways with six digits and asked to write them down from immediate memory, in any order they chose. Digits presented to the ear came from a Ferrograph tape recorder, carrying two channels on the same tape so that rigid synchronization was preserved. The head-phones employed were S. G. Brown Type D.1. Digits presented to the eye were typed with an Imperial typewriter, *élite* type, on a paper belt which the subject could see through a slit at his own comfortable reading distance. The belt could be moved suddenly, to bring a new portion of the paper into view, by a ratchet operated by the experimenter. The latter wore a head-phone and took the time for moving the paper from the recorded auditory presentation, even when the subject was receiving only visual signals. Thus the timing arrangements were similar for all conditions. This method of synchronization is, of course, crude, and restricts the results available from the experiment to those given below. Time measurements of the type given in Experiment II, for instance, would be quite unjustifiable.

Group I (12 subjects) received four different conditions, namely, straightforward auditory and visual memory spans as controls, a simultaneous auditory and visual presentation and a binaural presentation of the type used by Broadbent (1954). In the *auditory* condition all six digits were heard on the right ear only, spoken successively at the rate of two a second. In the *binaural* condition, two different digits were presented simultaneously, one to one ear, and one to the other. Another pair followed after half a

second, and then a third pair. In the *bisensory A* condition, two digits were presented simultaneously, one to the right ear and the other to the eye. Another pair followed after half a second, and then a third pair. In the *visual successive* condition all six digits were presented to the eye, successively, at the rate of two a second. Each subject received five lists under each of these conditions, half the individuals receiving them in the order described and half in the reverse order.

The visual presentation in Group I was somewhat artificial, as it is unnatural to watch digits rapidly appearing and disappearing, and a second group was therefore tested in which a more usual visual display was employed.

Group II (12 subjects) also received four different conditions. Two of these were visual successive and auditory, just as in Group I. A third, however, the *visual simultaneous*, consisted of presentation of all six digits together, typed in a line: they remained visible for the same length of time as that taken to present the whole series in the auditory and visual successive conditions, namely three seconds. The fourth was a *bisensory B* condition, in which three digits were delivered successively to the right ear. Three other digits were presented visually all together, appearing as the first auditory digit was spoken and disappearing as the last was completed.

Each subject received ten lists under each of these conditions; the experiment was spread over two days, on one of which the visual successive and simultaneous conditions were given, while on the other, the auditory and bisensory B condition were given. Half the subjects received each day first; within each day five lists were given under one condition, then ten under the other and finally, the remaining five of the first condition. All subjects were naval ratings, and none served in both groups.

TABLE I
PER CENT. CORRECT IN THE VISUAL-AUDITORY EXPERIMENT
(EXPERIMENT I)

Condition	Group I	Group II
Auditory	92	90
Visual Successive	60	43
Binaural	62	—
Bisensory A	77	—
Visual Simultaneous	—	78
Bisensory B	—	95

(a) Results.

The number of cases in which all six digits were correctly reproduced is given in Table I as a percentage of presentations for each condition. It is obvious that a considerable degree of success is achieved in the two bisensory conditions; our first conclusion is, therefore, that the eye and ear can to some extent receive different stimuli simultaneously.

Secondly, this adequate response to simultaneous stimuli is achieved by delaying all the information on one channel. When all six digits are correctly reproduced, it is normal for all three digits presented by one sense to be written down before any on the other channel. In only three cases out of 160 correct responses was any other order obtained.

We must therefore abandon the possible view that this phenomenon is peculiar to binaural hearing. The next point at issue is whether the delayed information is always that presented to the ear: is there a store which is associated only with the ear and not with other senses?

In fact this does not seem to be so, as there was no tendency for either sense to be universally the one delayed. In general each subject had his own consistent

preference for giving the eye or the ear first, but in each group only half the subjects gave the eye first.

A last point concerns the relative efficiency of bisensory and orthodox means of presentation. In the earlier paper it was argued that memory proper was as adequate in one case as in the other. Admittedly the binaural condition gave a lower proportion of successes than the auditory one, just as it did in Group I of the present experiment. (Seven subjects show the effect and none reverse it, so $p < 0.01$). But it was also argued from a control experiment in the previous paper that this loss of efficiency was not in the storage but in the intake of information. The present results support this argument, since there is no evidence that bisensory performance is any worse than orthodox presentation. In Group I, Bisensory A gives about the same as the mean of auditory and visual successive. It is worse than auditory alone (though not significantly) but this is probably because of the unnatural nature of the visual presentation. In Group II it is demonstrated that visual successive is really inferior to visual simultaneous; all subjects show the effect, so $p < 0.01$. Bisensory B is actually superior to the mean of auditory and visual simultaneous: this difference is significant ($t = 2.31$; at 0.05 level $t = 2.20$), but must be regarded with caution. It might be due, for example, to the artificial upper limit which is set on auditory performance by the highest possible score. At least, however, it is clear that the two senses together are certainly no worse than either alone, and the argument put forward previously for the binaural case is therefore supported: provided measures are taken to make the two sensory channels distinctive, small amounts of information can arrive simultaneously by two channels with as good a chance of being perceived as if they arrived successively.

It should also be noted that this high level of efficiency in the bisensory B condition can hardly be ascribed to alternation of attention, although the opportunities for such alternation seem greater than in bisensory A. All responses in this group gave one sensory channel completely before beginning the other; it is likely, from the evidence of the next experiment, that the time intervals allowed were too short to allow alternation.

III

EXPERIMENT II

HIGH AND LOW FREQUENCIES

Instead of using the two ears as separate sensory channels, it is possible to deliver two messages through the same loudspeaker, but to remove, say, all the low notes from one voice and all the high notes from the other. Since these experiments were begun Spieth, Curtis and Webster (1954) have published results showing that such a technique is useful in helping a listener to pick out one of two messages and ignore the other, just as natural differences in voice are (Broadbent, 1952). Voices may differ in this way by a greater or lesser extent, whereas the bisensory conditions we have dealt with so far are sharp and cannot be quantitatively varied. It is of interest to see whether the appearance of the effect we are considering depends on the degree of difference between the sensory channels, and this suggests the use of filtered voices. In addition, high and low frequencies produce activity at neighbouring points in the cortex, whereas the two ears gave points of maximal activity on opposite hemispheres and the eye and ear in widely separated projection areas. By suitable choice of a fixation point it might even have been possible for the subjects in the previous groups to bring the points of maximal activity due to the auditory and visual stimulation into different hemispheres. This hardly seems possible when different auditory frequencies

are used as the sensory channels, and a demonstration of the phenomenon with filtered voices thus adds to its generality. The same general type of experiment was therefore repeated with voices slightly and markedly distinguished by filtering.

A second point may also be investigated using filtered voices. In the earlier paper it was found that once a voice had been selected for response a certain length of time was required to shift to another channel. The technique adopted to establish this was to present sets of digits under the binaural condition, with instructions to reproduce them in the actual order of presentation, i.e. writing down the first digit from both ears before the second digit from either, and so on. This task was very difficult indeed when digits were presented at a $\frac{1}{2}$ -second rate or a 1-second rate, but easier at a $1\frac{1}{2}$ -second rate and possible for almost all subjects at a 2-second rate. The underlying argument of the experiment is that reproduction of material from immediate memory is much easier when required in the order in which it was heard than when required in any other order. Consequently when the effect of a particular order of reproduction changes as a function of the presentation conditions, it is reasonable to suppose that the material is being heard in a different order. The experiment quoted thus shows that a time of $1\frac{1}{2}$ –2 seconds is required for the analysis of a spoken digit, a shift to the other ear and analysis of a second digit, and a shift back again. But does this time vary when other sensory channels are used?

(a) *Method.*

One group of twelve naval rating subjects was employed, and received four conditions of which the first two are intended to show the effect of different degrees of filtering. In the *slight cut conditions*, the same tape recordings used for the binaural condition of Experiment I were mixed into a single loudspeaker. The subject thus heard a pair of digits spoken simultaneously through the speaker, followed half a second later by another pair, and then by a third. All digits were spoken by the same male voice, but before mixing, one channel was passed through a high-pass filter which attenuated the parts of the spectrum below 225 c.p.s. One digit in each pair, therefore, sounded slightly different in quality to the other. Instructions were to write down all six digits in any order the subject chose.

In the *sharp cut* condition, the high-pass filter was set at 1,280 cycles, all other features of the situation remaining the same. The two digits of each pair were thus more sharply distinguished in quality.

The other two conditions were concerned with comparing the time taken to shift attention with binaural and filtered presentations. In the *binaural shift* condition, the experiment of Broadbent (1954) was exactly repeated; binaural presentations of six digits, three to each ear, were given with instructions to write down the digits in the actual order of arrival. After a practice presentation three such presentations were then given at a rate of one pair of digits every 2 seconds, three at $1\frac{1}{2}$ seconds, three at 1 second, and three at $\frac{1}{2}$ second; the same four sub-conditions were then repeated in reverse order.

In the *spectral shift* condition, the same recordings were presented with the same instructions, but both channels were mixed into a single loudspeaker. One channel was passed through the high-pass filter at the 1,280 c.p.s. setting before mixing, so that the two digits in each pair differed in their spectrum.

The subjects were tested each on two successive days. To rotate the order of receiving the conditions, the group was divided into four sub-groups of three subjects each. Sub-group A received on the first day five presentations of the type labelled auditory in Experiment I, followed by five under the sharp cut condition, then five under the slight cut, and another five auditory presentations. They then received the binaural shift condition. On the second day, they repeated this sequence with the order of the sharp cut and slight cut conditions interchanged, and with the spectral shift and spectral shift conditions interchanged. Sub-group B received the same sequence as A, but with the slight and sharp cut conditions interchanged throughout, and sub-group D the same sequence as C with the binaural and spectral shift conditions interchanged. The loudspeaker was a Grampian G. 129: the filters were designed and built by Dr. A. Carpenter and had an attenuation of at least 18 db per octave.

TABLE II

PER CENT. CORRECT WITH VOICES DISTINGUISHED BY SPECTRAL FILTERING

	Whole Experiment		First Condition of each Subject	
One voice bass cut at	1,280 c.p.s.	225	1,280	225
Digits mixed up in response ..	13	19	3	20
Two voices distinct in responses..	46	37	36	17
Auditory for same subjects	98			

TABLE III

PER CENT. CORRECT AT VARIOUS RATES OF PRESENTATION WHEN SHIFTING ATTENTION BETWEEN TWO CHANNELS

Presentation ..	Rate			
	$\frac{1}{2}$	1	$1\frac{1}{2}$	2
Binaural	14	25	37	54
Spectral	19	33	54	47

(b) Results.

We may consider first the effect of varying the amount of cut, which gives some support to the view that one channel becomes more likely to be delayed as the difference between the channels increases. The percentage of correct responses is shown in Table II, separate data being given for responses in which all the material from one channel was written down before any from the other. It will be noted that the total number of occasions on which all six digits were correctly recorded is much the same with either degree of filtering, but that there appears to be a slight tendency for the sharp cut to give a higher proportion of these correct responses in the order already found for binaural and bisensory presentation. The pooled results for the entire experiment are difficult to interpret, as on analysis of variance one of the degrees of freedom between the sub-groups of subjects was found to be significant: apparently subjects who adopt one mode of performance on their initial condition may then apply it to subsequent conditions.

Consequently Table II also gives the results for subjects on their first five runs (ignoring auditory runs) where there are no effects carried over from previous conditions. In this case, the difference becomes even more striking: all the subjects of the sharp cut condition show less responses given in a mixed order from the two channels, so $p < 0.05$. Only one of the slight cut subjects shows this type of performance, giving a total difference insignificantly in the opposite direction. The difference between the two conditions is significant, $p < 0.05$, on constructing a 2×2 table for the subjects showing each type of performance.

Thus on naïve subjects and with two voices sharply distinguished by filtering out certain frequencies, the now familiar phenomenon seems to occur; while with less marked filtering some alternation or confusion of the two channels may occur. But

transfer effects are not clear and so experienced subjects may not show these differences. This moderate effect is somewhat to be expected if successive response to two channels requires the channels to be separate: even the sharply-filtered voice shares more sense-organs with its partner than does a voice on one ear with a voice on the other ear. However, it is clear that our phenomenon may appear with filtered voices and so is probably not an effect depending on the distinction of the two hemispheres.

On our second point of interest, the results on the time taken for shifting attention are shown in Table III: they indicate that shifts between two voices distinguished by filtering can be made more rapidly than can shifts between the two ears. As in the previous paper, efficiency under the binaural condition increases as the rate of presentation slows down, reaching at 2-seconds roughly the level attainable when the order of reproduction is not controlled. The same improvement appears in the spectral shift condition, the differences between the 2-second rate and the $\frac{1}{2}$ -second rate being significant ($t = 2.38$; at 0.05 level, $t = 2.20$). But with the spectral shift the performance at fast rates is more effective than with the binaural shift: for the three faster rates together $t = 2.23$. The 50 per cent. level (which corresponds roughly to the performance attainable under uncontrolled order of reproduction) is reached at $1\frac{1}{2}$ seconds instead of 2 seconds.

IV

DISCUSSION

It is probable from the results of these experiments that temporary storage of information in one sensory channel is possible with a number of different senses, and certainly it is not peculiar to the binaural situation. It may, however, be replaced by alternation of attention or confusion of the two channels if the sensory cues associated with the two channels are not very distinctive. It seems likely that liability to confusion is associated with the possibility of rapid shifting of attention, so that different methods of presenting messages to a listener may be appropriate for different tasks. Two messages distinguished by filtering may allow the listener to alternate rapidly, but confusion is more common than with a spatial separation of messages. A bisensory distinction, one message to the eye and one to the ear, is the most effective in these results, presumably because of the lack of perceptual confusion: a notable omission in the present data is the measurement of the time taken to shift attention between the ear and eye, which was beyond the scope of the apparatus available. It is hoped to determine this value at some later date.

Quite apart from the minor changes which may be produced by altering the sensory channels used, however, this mode of response to simultaneous stimuli seems to be a general one. Mowbray was perfectly correct in emphasizing that some processes in the nervous system are incompatible and cannot occur simultaneously; his experiments isolated such processes, both by using material of high information context and also by inserting neutral irrelevant stimuli on the same sensory channel before and after his relevant stimuli. (The latter factor has been shown to be important in experiments as yet unpublished.) But in addition to these mutually incompatible processes there are others which may occur simultaneously, and some of which must in the present case be regarded as lying earlier in the chain of events between stimulus and response. The information enters simultaneously; it leaves successively. Such a process implies a storage or delay system, but it remains to be shown how this delay system is related to more familiar problems of immediate memory.

This work was done under the general direction of Dr. N. H. Mackworth, and thanks are due to the Royal Navy both for supplying subjects and providing equipment.

REFERENCES

- BROADBENT, D. E. (1952). Failures of attention in selective listening. *J. exp. Psychol.*, **44**, 428-433.
- BROADBENT, D. E. (1954). The rôle of auditory localization in attention and memory span. *J. exp. Psychol.*, **47**, 191-196.
- HEBB, D. O. (1949). *The Organization of Behaviour*. London.
- MOWBRAY, G. H. (1953). Simultaneous vision and audition: the comprehension of prose passages with varying levels of difficulty. *J. exp. Psychol.*, **46**, 365-372.
- MOWBRAY, G. H. (1954). The perception of short phrases presented simultaneously for visual and auditory reception. *Quart. J. exp. Psychol.*, **6**, 86-92.
- SPIETH, W., CURTIS, J. F. and WEBSTER, J. C. (1954). Responding to one of two simultaneous messages. *J. acoust. soc. Amer.*, **26**, 391-396.

DIFFERENT THRESHOLDS FOR RECOGNITION— FURTHER EXPERIMENTS ON INTERPOLATED RECALL AND RECOGNITION

BY

HARRY KAY and RICHARD SKEMP

From the Institute of Experimental Psychology, Oxford University

Three experiments are described which have attempted to identify some of the difficulties in recognition after interpolated recall. In the first, complex picture material was used, and, though subjects were able to recognize the original after interpolated recall, they did in fact recognize fewer items. This posed the query, why should the recall of some items depress the recognition of other unrecalled items? The second experiment showed that in recognition the stronger (correct) memory has an inhibitory effect on a weaker (also correct) memory. The third experiment therefore examined the hypothesis that the juxtaposition of better and worse known items raises the threshold of recognition of the worse known, and confirmed that the threshold of recognition for an item varies according to its context. The theoretical significance of this finding is briefly indicated.

INTRODUCTION

Experiments upon the differences and interaction between recall and recognition have a long history, and have aroused both methodological and theoretical considerations. The following group of experiments was prompted partly because it was felt that the experimental evidence to date may be "the truth but not the whole truth," and partly because certain theoretical issues can be neatly scrutinized when the facts are presented within this framework of initial perception, interpolated recall and subsequent recognition. The first two experiments follow this pattern, and the third attempts to isolate a variable suggested by their results.

From previous work in this field three different approaches may be distinguished. The first, stemming from Ebbinghaus, has attempted quantification of recall and recognition and has been interested in methodological issues. Hollingworth's paper (1913), Luh's work (1922), and more recently Postman's (1948) are in this tradition—they present such findings as that much more material is recognized than recalled, and that items which are recalled are generally recognized, though the converse does not hold that recognized items tend to be recalled. A second approach is represented by Gestalt theory and its postulation of progressive autonomous changes in the memory trace. This work, culminating in Koffka's (1935) contribution, is experimentally examined by Zangwill (1937) and effectively criticized by Hanawalt (1937) and Hebb and Foord (1945). Thirdly there is a large and interesting group of experiments, which might be said to represent "Variations on a Theme by Bartlett"—these are qualitative studies on the merging of experiences (e.g. Davis and Sinha, 1950), and the similarity of various influences upon both recognition and recall. It was an experiment in this third group by Belbin (1950) which was the starting point for the present studies.

Experiment I

Postman (1948) has shown that, as the layman presumes, recall is better after previous recognition, and also, as the layman does not presume, that recognition

tends to be worse after recall. He attempted to explain this on a "minimal strength of association" hypothesis, such strengths differing for recognition and recall. Belbin's experiment was a striking illustration of this influence of recall upon recognition. Of twelve subjects who were shown a newspaper picture with one central incident and were then asked to recall it, only two could subsequently recognize it. Of twelve other subjects who were not asked to recall it, ten correctly recognized it. The author discussed her findings on Bartlettian lines—recall introduced erroneous importations, related experiences were responsible for the insertion of errors, etc. In preliminary studies we used similar news illustrations to repeat the experiment. The results with five groups of subjects—in all over sixty subjects—tended to confirm Belbin's results. None of our results was so striking—the tendency was for half the experimental group to recognize and half not to do so after any interpolated recall, but this compared strikingly enough against control groups in which few had any difficulty in recognizing the original. The features noted by Belbin were influential, but there were others which brought out the complexity of the processes. For example, as noted by Claparède (1911) and by Koffka (1935) it was not infrequent for the recall of some part of the illustration to be wide of the mark and yet for subsequent recognition to be correct. One of our subjects recalled a picture of a cattle show, with tractors in it, as a picture of a motor show. When later confronted with the original he recognized it instantly, agreed he had recalled it as a motor show, but commented, "You cannot fool me now."

Belbin had concluded that it was likely that "the amount of correct detail which a subject is able to recall is less important in determining eventual identification of the original material than the number of items he erroneously imports in his effort." Certainly importations played their rôle in her experiments as it had in previous work (e.g. Zangwill, 1939), but on the other hand both in our pilot studies and in Belbin's the experimental picture material was concerned with one central incident. It was possible that this kind of material favoured the above conclusion and it was therefore decided to carry out an experiment with material where many incidents were depicted. Picture material was used rather than a set of disconnected items such as nonsense syllables since we wished to know how far a subject might accept or reject material as a whole while perhaps making some reservations about the constituent individual items.

Method

The picture was taken from the 1948 cover of a children's magazine. It was a water colour of a London park scene, depicting a lake, lawns and pathways with groups of adults, children, rowers, and horsemen. Two copies of the picture were mounted to conceal all names and show an area 7 in. $\times$ 8 in. The experimenters listed 30 items from the picture which were to serve as a recognition test; this list covered all details which a pilot group of subjects had been able to recognize immediately after presentation.

The experimental design was as follows:—

Experimental Group ..	Presentation of Picture (40 seconds exposure).	Written recall (10 minutes later).	Recognition test (40 minutes later) i. of picture, ii. 30 item ques- tionnaire.
Control Group	Ditto	—	Ditto

In the recognition test each subject was shown the original and asked to say whether it was or was not the same as he had been shown previously. After deciding, each was handed a questionnaire in which the 30 items were listed and he was asked to say whether he recognized them. He identified these items from a key illustration placed alongside the same picture as used in the recognition test. He was not told whether this picture was the

original. After filling in the questionnaire he was asked to write any comments upon his judgments and whether he wished to change his original decision.

Subjects

There were 21 subjects in the experimental group and 15 in the control. They were made up of university students, and students attending a W.E.A. Summer School, whose age ranges were from 20 to 60. There were slightly more men in each group than women.

During the intervals between presentation and recall and recognition all subjects were occupied in viewing a rotating trapezoid and making judgments upon it. They were fully engaged during the period and were not aware that it was not part of the actual experiment.

Recall

RESULTS

Table I shows that the experimental group recalled in all 153 items out of a possible 630, the mean being 7.3 items (approximately 24 per cent.), and the S.D. being 2.45. Two points are of note. Firstly, in spite of seeing the picture for a relatively long period—40 seconds—the recalled items were only a small proportion of the possible total. It did not seem that subjects had a clear memory of many separate incidents, but a general impression of fun and gaiety. They all mentioned that it was a park scene and many elaborated on the Saturday afternoon or Bank Holiday atmosphere of it.

TABLE I
RECALL SCORES OF EXPERIMENTAL GROUP. (21 subjects.)

CORRECT RECALL:					
Total number of items (possible total 630)					153 (24.3 per cent.)
Mean number per subject					7.28
S.D.					2.45
INCORRECT RECALL:					
Total number of items					41
Number of subjects making no errors ..					2

Secondly, recall of individual items was by no means accurate. In all there was a total of 41 wrong items, of which 20 were inaccurate memories of actual features in the picture and 21 were importations of events which might have been but were not present. All except two of the 21 subjects had some inaccuracies, of which many were striking—"a haystack" for the willow tree; no females in the picture where there were eight; two children on ponies for one; boys playing with a ball when there were none. One subject alone mentioned "six men carrying a pole" (for four men carrying a boat), "a large man uncomfortably asleep on an iron bedstead" (the tramp on the park bench) and "a man selling something from a barrow to a queue of laughing children" (an importation, but possibly inspired by the artist and his easel with two children looking at it). If erroneous recall does result in subsequent failure to recognize material many of these subjects might have been expected to reject the original.

Recognition

At the recognition test of the 21 experimental subjects two rejected the original; of the 15 controls every one affirmed it was the same. Both subjects who failed to recognize the original insisted that it was very similar ("90 per cent. the same") but that some details were changed. On the face of it this result differs considerably from those of our own pilot studies

and of Belbin's. Here, with complex material, though only a proportion of items could be recalled correctly, there was still a sufficiency of features, either general or individual, for the original to be recognized. One striking difference between this study and our pilot experiments with simpler material was that subjects obviously did not expect to be able to recall or recognize all of the complex material. They had lowered their standard. The experimental group was more prepared to accept details which had not been recalled. The larger number of items reduced the contribution of each to the whole, both in the initial perceiving and in subsequent remembering, and lowered the significance of an importation in recall. It appeared that a subject was prepared to say it was the original when there was a sufficiency of recognizable items which together conveyed an overall feeling of familiarity. This general impression was obviously important to subjects, as we noted of the written recall. Subject No. 7 commented in his report after the recognition—"I do not think any one of these details was particularly important, but they all added up to certainty. I also recognized the statue and the kite which I had forgotten at the time of recall. But I was surprised that the rowing crew carrying boat, were carrying a boat. I remember them just as bathers with pink skins, and not with striped shirts. Had there been more surprises of this sort I might have been less certain." This comment exemplifies the point; the subject had in fact recalled the rowing crew carrying their boat as "a row of pink-skinned people on a sort of diving board," but was prepared to reject this striking memory because of the wealth of contrary evidence.

Recognition of individual items

Subjects were presented with the list of 30 items and, with the picture in front of them, asked to score how many they recognized. They did not know whether it was the picture which they had originally seen. Table II shows that there was a

TABLE II
RECOGNITION SCORES

	<i>Experimental Group</i> (21 subjects)		<i>Control Group</i> (15 subjects)	
	per cent.		per cent.	
Recognition of 30 items (mean) 	43.5		64.2	
Recognition of recalled items or equivalent 10 items for control group 	88.2		80.6	
Recognition of non-recalled items or equivalent 20 items for control group 	29.1		56.0	
Recognition of picture as whole 	Yes 19	No 2	Yes 15	No —

marked difference between the number of recognizable items in the two groups, only 43.5 per cent. for the experimental as against 64.2 per cent. in the control. These figures seemed worthy of further analysis. Of the 153 recalled items in the experimental group 135 (88.2 per cent.) were recognized. In order to make a comparison with the control group where there was no recall, the recognition scores were examined and the ten most popular items selected. As would be supposed, some of the more obvious items such as the path, the lake, the tramp asleep on the bench, etc., were recalled much more frequently than others. All the ten selected were recalled by

more than half of the subjects in the experimental group. On these items the control group also scored highly—121 out of 150 (80.6 per cent.). The interesting question was what happened in recognition to the other 20 less popular items. We find of these that the control group recognized 168 out of 300 (56 per cent.) whilst the experimental group over similar items which were not recalled scored 139 out of 477 (29.1 per cent.). It would appear then that whilst the experimental group scored highly enough on recalled items they recognized few which they had failed to recall, and in fact significantly fewer than did a control group over approximately identical items. The question to be answered was, why should the recalling of some items depress the recognition of others unrecalled? This was obviously a question of some theoretical importance, and the second and third experiments were designed to offer some solution to it. Discussion will be deferred until after the results of these experiments.

Experiment II

INTRODUCTION

This and the following experiment had their beginning in an attempt (here briefly reported) to investigate the effect of an increasing number of errors in the interpolated recall upon subsequent recognition. Outline faces were shown in which variations of four main details were permuted: the ears being flat or round, the eyebrows arched or straight, the nose formed of two dots or of two lines, and the mouth in five different shapes. Ten such "faces" were shown briefly to the subjects (27 school-boys aged 12-14). Immediately after the individual presentation they were asked to draw five of the faces from memory; the other five they merely observed. Without interval (in view of the difficulty of the material) there followed the recognition test. The ten originally presented faces were shown randomly mixed with ten new ones, and the subjects asked to say if they recognized them. Chance guessing would have given a score of 50 per cent. In fact they scored as follows:—

TABLE III

<i>Pictures observed and—</i>	<i>Percentage correctly recognized</i>
not recalled	56
correctly recalled	63
recalled with 1 error	56
recalled with 2 errors	39
recalled with 3 or 4 errors	25

It seemed clear that recall with more than one error had depressed the recognition score below the chance level. The effect could have been explained in more ways than one. Interpreting it along the lines of Bartlett's schema, it would appear that the original memory had itself been modified by recall and that the resulting inaccurate version was thereafter the only available criterion for recognition. Another possibility was that the original memory coexisted with that of the interpolated recall, but the latter, being more recent and having involved greater activity by the subject, had a lower threshold. It was therefore used as criterion. Under such a hypothesis, it would seem probable that in favourable circumstances the earlier memory could be revived, and the following experiment was designed to test this.

METHOD

Four lists, A, B, X, Y, were prepared, each containing 15 familiar names. Lists A and B were intermingled in a brief narrative which was read out twice to the subjects. Immediately after this presentation they were asked to write down as many of the names as they could recall. These "recalled lists" were then collected.

After an occupied interval two consecutive recognition tests were performed. Each used half the originally presented names (i.e. only list A or B) randomly interspersed with an equal number of new names (list X or list Y). For the first recognition test, hereafter called "isolated," only the mixed original and new names (i.e. list AX or BY) were read to the subjects. For the second, hereafter called "in narrative," the names were presented in the narrative with the originals (either list A or B) placed in their correct context and the new names (list X or Y) replacing those from the omitted list.

All this was done initially with two groups of subjects, using for one (group Ia) a 10-minute and for the other (group IIa) a 25-minute interval. This design made it possible to investigate the effects of a second variable which might affect the use of the original impression as criterion, namely length of interval between recall and recognition in addition to restoration of narrative. It will be seen that with this design each group was acting as its own control.

When both these factors were present, the results were sufficiently striking to make it seem desirable to confirm them by a repetition of the experiment in exactly the same form with another group of subjects. These are here called Ib (10-minute interval) and IIb (25-minute). The two sets of results showed good agreement, and are given in Table IV, together with the combined scores of both sets of subjects. Group Ia and Ib are combined as group I, and groups IIa and IIb as group II. In view of this close agreement, we may hereafter confine our attention to the combined scores. A chance score would have been 50 per cent.

Subjects

The first set (groups Ia and IIa) consisted of 27 schoolgirls aged 12-14. The second set (groups Ib and IIb) were 30 schoolboys aged 10-13 at a different school. None knew the purpose of the experiment.

RESULTS
TABLE IV

	Subjects					
	1st set		2nd set		Combined	
	Ia (10-min. interval)	IIa (25-min. interval)	Ib (10-min. interval)	IIb (25-min. interval)	I (10-min. interval)	II (25-min. interval)
Recognition when isolated:	per cent.	per cent.	per cent.	per cent.	per cent.	per cent.
of those recalled ..	100	98	98	93	99	95
of those not recalled ..	64	62	65	65	65	63
Recognition in narrative:						
of those recalled ..	99	96	97	95	98	95
of those not recalled ..	57	91	61	82	60	86

Significance.—For the non-recalled names in the narrative the difference between the 60 and 86 per cent. scores is significant at the 0.01 level of confidence.

Discussion

The interest in these results centres around the non-recalled names; i.e. those initially presented to subjects but omitted by them from the recalled list. For it is improved recognition of these which we should take to indicate revival of the original impression.

It will be seen from Table IV that for "isolated" recognition items which had not been recalled were recognized approximately the same with the 10- and 25-minute intervals (65 to 63 per cent.). This result agreed well enough with those of previous experimenters. After 10 minutes the recognition scores were, if anything, better for

those names presented in isolation than for those in narrative. But unlike the isolated names there was considerable difference between the recognition scores for the names in the narrative after 10- and 25-minute interval—60 per cent. recognition after 10 minutes and 86 per cent. after 25-minute interval. Thus 26 per cent. of the non-recalled names could not be recognized with a 10-minute interval, yet were recognized after one of 25 minutes. It had now been shown to be possible, as the second hypothesis had supposed, to revive the original memory. That this took place with the longer and not the shorter interval was surprising, but further supported the hypothesis. The "reminiscence" effect could only mean that such memories were never wholly absent, but that some factors or factor made them in the first instance temporarily unavailable—a factor operating more strongly with the shorter interval.

If the obvious line of argument is now followed, that the interfering factor was the interpolated recall and that this was relatively weaker after the longer interval, there is a further problem. We cannot in this experiment suppose interference by incorrect recall, as was possible with the experiment outlined in the introduction and in earlier work (e.g. Belbin, 1950): for the recalled list, which we are supposing to have been interfering with recognition of certain names, did not contain these names. One thing seems clear, that whereas the effect of an erroneous interpolated recall does admit of an explanation by the hypothesis that the original memory was thereby modified, this new effect does not.

Accepting therefore, as a working hypothesis, that the effect indicated by the comparative scores for recognition of non-recalled names (an increase from 60 per cent. with the shorter to 86 per cent. with the longer interval) may be explained as a diminution of interference by the interpolated recall, a further question remains. Does the presence of a strong (correct) "trace" have an inhibitory effect on a weaker (also correct) one, as this experiment seems to indicate? It will be remembered that the results of Experiment I also pointed in this direction. Experiment 3 was a first attempt to isolate and confirm this phenomenon.

Experiment 3

INTRODUCTION

The hypothesis was formulated as follows:—"That the juxtaposition of better and worse known items will lower the probability of recognizing the latter." It was therefore desired to give an experimental group a higher certainty on some items of a list, the remaining items being equally presented to this and a control group. On the experimental hypothesis, the E group should score worse than the C group in a recognition test of these remaining items.

METHOD

The material used for the initial impression was a list AB of 20 familiar names, made up of two lists A (5 names) and B (15) intermingled. This was read aloud twice to the subjects. Immediately before the reading, the E group was given list A written on paper and asked to learn it by heart. Otherwise both groups were treated alike and were taken together. The E group knew that the pre-learned names would be part of the main learning task. After a 10-minute interval list AB was presented for recognition intermixed with 45 new names (list C). The subjects were told that some or perhaps all of the names in the list first read out would be given, together with some new ones. The experimental design was put ticks or crosses to indicate recognition or otherwise. The experimental design was thus—

C group

presentation of
list AB

After 10-minute
interval
Recognition test
with list ABC

E group

prelearn list A

ditto
Bureau Ednl. Psy. Research
DAVID HA - LALING COLLEGE

There was no interpolated activity for either group. The main learning task was somewhat easier for the E group, since they had only 15 new items to learn instead of 20. This factor would operate against the effect predicted by the experimental hypothesis, namely that E group would score worse than the C group on list B as a result of the presence of the better known items of list A. Thus any result in the direction indicated would support the hypothesis.

Subjects

The experiment was first conducted with 10 adults from a W.E.A. summer school and subsequently with 28 schoolboys aged 12-16. No subject had been in any of the previous experiments.

RESULTS
TABLE V
RESULTS OF RECOGNITION TESTS

Lists	CORRECT				INCORRECT
	A	B	C	B and C combined	C
	per cent.	per cent.	per cent.	per cent.	per cent.
10 Adult Subjects:					
Control Group	92	88	88	88	12
Experimental Group ..	100	76	74	74	26
28 Schoolboy Subjects:					
Control Group	66	86	92	90	8
Experimental Group ..	99	82	80	80	20

NOTE: Correct answers to Lists A and B = items recognized.

Correct answers to Lists C = items correctly recognized as not being present at initial presentation.

Incorrect answers to List C = items mistakenly recognized as being present at initial presentation.

At first glance, these results do not show a very big difference in scores on list B, though it is in the predicted direction. But what is very noticeable is that the E group mistakenly "recognized" a much larger number of names in list C: with the adult subjects, 26 per cent. as against 12 per cent. by the control group, and with the schoolboys, 20 per cent. as against 8 per cent. Both these differences are statistically significant at the 0.01 confidence level. Thus over much of the material there was greater uncertainty about recognition.

It will be appreciated that any increase in the total number of "Yes" answers would increase a subject's score on list B. Since the experimental group was answering "Yes" more frequently to list C, it is probable that their scores on list B contain a larger proportion of apparently correct recognitions which were really due to this factor. On the assumption that the proportion of guesses was the same for both lists, we can correct the list B scores for both E and C groups: reducing these scores by a percentage of themselves equal to the percentage of (incorrect) "Yes" answers to list C by that group. If we do this, we obtain the following corrected recognition scores for list B:—

Adult subjects: C Group, 77 per cent. E Group, 56 per cent.

Schoolboy subjects: C Group, 79 per cent. E Group, 66 per cent.

It may thus be argued that the results support the experimental hypothesis, showing in the raw scores as an increase in inaccurate guesses or in the corrected

scores as a decrease in presumed genuine recognition scores (as opposed to correct guesses) by the E group.

Discussion

The experiment would seem to have successfully isolated the phenomenon which was inferred from Experiments 1 and 2, namely that the presence of well-known items depresses recognition of lesser known items when these are juxtaposed.

It would appear that this feature was contributing to many of the studies of interpolated recall and recognition. Items which had been strengthened by recall tended to depress the recognition score on items not so recalled.

It is, however, realized that there are possible objections to this experiment. One is, that the pre-learning of list A by the E group may have affected the learning task itself for them. They might for example have been so busy looking out for the names they had already learnt that they paid less attention to the others. It might, of course, equally well be argued that they were more likely to concentrate on learning the new ones, and to pay less attention to those pre-learned. But the pre-learning task does give rise to an uncertain effect; and amongst further experiments which are planned is one to avoid this feature, by using material in which some items will initially be better known than others and so avoid the necessity for either previous or interpolated activity.

GENERAL DISCUSSION

Here it is proposed to discuss very briefly the implications of the findings. Accepting the general trend which runs through all three experiments, we have established that there is a variable threshold for recognizing the same item—in these instances the same items were or were not recognized according to their context.

The finding adds to the experimental evidence upon a relatively narrow feature of remembering. *This is all to the good since it would seem that in such a broad field as remembering intensive research on particular issues is now likely to be most productive. To date, an outstanding difficulty in considering available evidence is that much of it is compatible with different theoretical interpretations. It is suggested that the present finding isolates a phenomenon which is consistent with relatively few. For example we are not required to consider a system which loses data and then recreates them, but rather one in which different data are competing with one another. Without becoming involved in the intricacies of how data are coded or classified we might make the modest postulate that a critical variable in remembering is the different strengths of the retained data. Such strengths would seem to be determined by variables which previous researches have noted, particularly those discussed by Bartlett. We might say that in a situation where a subject is trying to recognize a relatively unknown item in a context of many better known items, he is trying to detect a weak signal amongst many stronger signals. The "noise" of the channel defeats the detecting mechanism. But the signal itself is present and may be sufficiently above threshold value to be detected in a context where stronger signals are absent. Our hypothesis for Experiment 3 on the decreasing probability of recognizing items might be rephrased to read, "the juxtaposition of better and worse known items will raise the threshold of recognition of the latter."*

Phenomenally then an item may have been unrecognized, but the system itself has not lost it, though it has failed to detect or identify it on a particular context. This raises a further point. Perhaps because of the inherent difficulty involved in any theory of remembering there has always been a tendency for theories to be a restatement about the phenomena which they were supposed to be explaining. For

example in Gestalt psychology there is the phenomenon of closure from which is postulated an electrical brain field exhibiting similar closure characteristics. But systems need not be mere reduplications of the phenomenal features which they are trying to explain. In a very simple system motor-cars run along the ground but the pistons go up and down—there are no wheels to generate the power. In the field of remembering the present data may be of value in suggesting that by making our systems imitate too closely the phenomena they have to explain, we have created difficulties for ourselves. As noted above, we need not postulate a mechanism, which loses data only to be able to find them again. Rather we are studying the conditions in a system which make data more readily available on some occasions than on others.

It might be mentioned in conclusion that our findings are not incompatible with more general evidence, such as a learner continuing to repeat mistakes long after a time when he might have been expected to have learned the correct response. On the present hypothesis it would be granted that he has "learned" the correct response in the sense of having stored it in the system, but the very much greater strength of the error competes successfully against it. Learning is accomplished only when the strengths of correct and error responses become relatively equated. Whether such signals are merely competing against one another in a limited capacity channel or actively inhibiting each other would seem to be a matter for further research.

ACKNOWLEDGEMENT

The authors wish to thank for their co-operation the headmasters and headmistress, staff and pupils of the following schools:—Magdalen College School, The Dragon School, and Rye St. Antony, Oxford.

REFERENCES

- BELBIN, E. (1950). The influence of interpolated recall upon recognition. *Quart. J. exp. Psychol.*, **2**, 163-169.
- CLAPARÈDE, E. (1911). Recognition et Moi-même. *Arch. Psychol.*, **11**.
- DAVIS, D. R., and SINHA, D. (1950). The effect of one experience upon the recall of another. *Quart. J. exp. Psychol.*, **2**, 43-52.
- HEBB, D. O., and FOORD, E. N. (1945). Errors of visual recognition and the nature of the trace. *J. exp. Psychol.*, **35**, 335-348.
- HOLLINGWORTH, H. L. (1913). Characteristic differences between recall and recognition. *Amer. J. Psychol.*, **24**, 532-544.
- KOFFKA, K. (1935). *Principles of Gestalt Psychology*. London.
- LUH, C. W. (1922). The Conditions of Retention. *Psychol. Monogr.*, **31**, No. 142.
- POSTMAN, L., JENKINS, W. O., and POSTMAN, D. L. (1948). "An experimental comparison of active recall and recognition. *Amer. J. Psychol.*, **61**, 511-519.
- ZANGWILL, O. L. (1937). An investigation into the relationship between processes of reproducing and recognizing simple figures with special reference to Koffka's Trace Theory. *Brit. J. Psychol.*, **27**, 250-76.
- ZANGWILL, O. L. (1939). Some relations between reproducing and recognizing prose material. *Brit. J. Psychol.*, **29**, 370-382.

NON-RANDOM SEQUENCES IN VISUAL THRESHOLD EXPERIMENTS

BY

C. I. HOWARTH and M. G. BULMER

From the R.A.F. Institute of Aviation Medicine, Farnborough, Hants.

Responses (i.e. "seen" or "not seen") to a repeated stimulus of constant intensity are shown to be grouped in runs of the same response. It is suggested that there are only two possible explanations of this non-randomness: spontaneous fluctuations of threshold and the direct influence of a response on succeeding responses. Experiments designed to distinguish between these two explanations have shown that the latter is the more important and that a response affects directly, though with diminishing intensity, the three immediately following responses.

I

INTRODUCTION

The assumption that successive determinations of a visual threshold are independent of each other has been questioned in two recent papers (Verplank, Collier and Cotton, 1952; Wertheimer, 1953). They have shown that responses to successive stimuli of the same intensity tend to repeat themselves and form runs such as YYYYYY NNNNN . . . (Y stands for a stimulus seen, N for a stimulus not seen) rather than to alternate like YNYYNYN The purpose of this paper is to study in greater detail the form of these non-random sequences and to discover their cause.

The first question has been attacked by considering the conditional probabilities after particular patterns of response. For example the conditional probability of N after NY is written $P(N|NY)$ and is the proportion of times that N follows NY out of all the times NY occurs in the experiment. Obviously if runs preponderate over alternations:

$$P(N|N) > P(N|Y).$$

The question we have asked is: How far back does this relationship extend? Given that $P(N|N) > P(N|Y)$, is $P(N|NN)$ for example greater than $P(N|YN)$, and, if so, is $P(N|NNN)$ greater than $P(N|YNN)$ and by how much?

The frequent occurrence of runs of the same response indicates fluctuations of the visual threshold; this follows from the definition of the threshold. These fluctuations of threshold must either be spontaneous, due perhaps to some physiological rhythm, or be induced by the direct effect of one response on succeeding responses. The only other possible explanation is that there are fluctuations of the stimulus intensity; but these have been eliminated in our experiments. Spontaneous fluctuations of the threshold could produce an excess of runs by causing different probabilities of seeing in different parts of the experiment; then runs of "noes" would be more common in some parts of it and runs of "yeses" in others. On the other hand, a direct influence of one response on succeeding ones would produce runs if the fact that the subject had said "yes" once predisposed him to say "yes" again. In the first case the fluctuations of threshold cause the runs; in the second the runs are self-induced and cause the fluctuations of threshold. The second experiment reported in this paper was designed to distinguish between these two explanations.

II

APPARATUS

A 12-volt, 36-watt head-lamp battery provided the light source. The voltage supply was a 24-volt, 100-amp./hr. battery dropped through a series resistance. The voltage across the lamp could be varied through a small range by means of a rheostat in parallel with the series resistance. The voltage across the bulb was measured by comparison with a reference voltage and could be accurately controlled to within 0.025 of a volt. A projection system projected a $\frac{1}{2}$ -inch square of light on to a rear projection screen when a solenoid-operated shutter was opened. The shutter gave $\frac{1}{2}$ -second flashes separated by intervals of 4 seconds. A faint black cross acted as a fixation point. A small bell rang simultaneously with each flash. The screen had a background illumination of 10 millilamberts from a series of four 6-volt bulbs connected across the 24-volt supply. The intensity of the light flash was coarsely controlled by neutral-density gelatine filters. A small range of final adjustment was provided by the rheostat in parallel with the bulb.

A photocell connected with a valve bridge circuit was used to give a very sensitive monitoring of the intensity of the light. The intensity remained constant to within 1 per cent. over long periods. In a few experiments a slight correction was applied during the course of the experiment by means of the rheostat. The subject responded by pressing a morse key. Stimulus and response were recorded on the same pen on a Palmer drum.

III

RESULTS

Experiment 1: The form of the non-random sequences

Procedure. Ten subjects were used, 9 male, 1 female. They were all under the age of 30 and had no gross visual defects. Each was placed in the dark room 6 feet in front of the illuminated screen and allowed a few minutes to adapt. The apparatus was started and the light flashed every 4 seconds. The intensity of the light was adjusted as quickly as possible until the subject could see the flash on about half the number of times at a given intensity. A further 10 minutes was used on final adjustment of the intensity from outside the dark room. This gave the subject time to settle down and prevented any large deviations from 50 per cent. seeing during the main part of the experiment. Blank stimuli were introduced during this preliminary session, and the subject was warned if he gave too many false alarms. When all adjustment was completed, the subject was told that the period of adjustment was completed and that the main experiment would now begin. He was told to press the morse key when he saw the flash. No other instructions were given apart from warnings not to fixate too strongly between flashes and to rest the eyes by blinking. The subject then received about 600 successive stimuli at the same intensity. The experiment was stopped on two occasions after the probability of seeing had dropped almost to zero. This change was thought to be due to fatigue and both these subjects (O. B. and O. T.) performed satisfactorily on a later occasion.

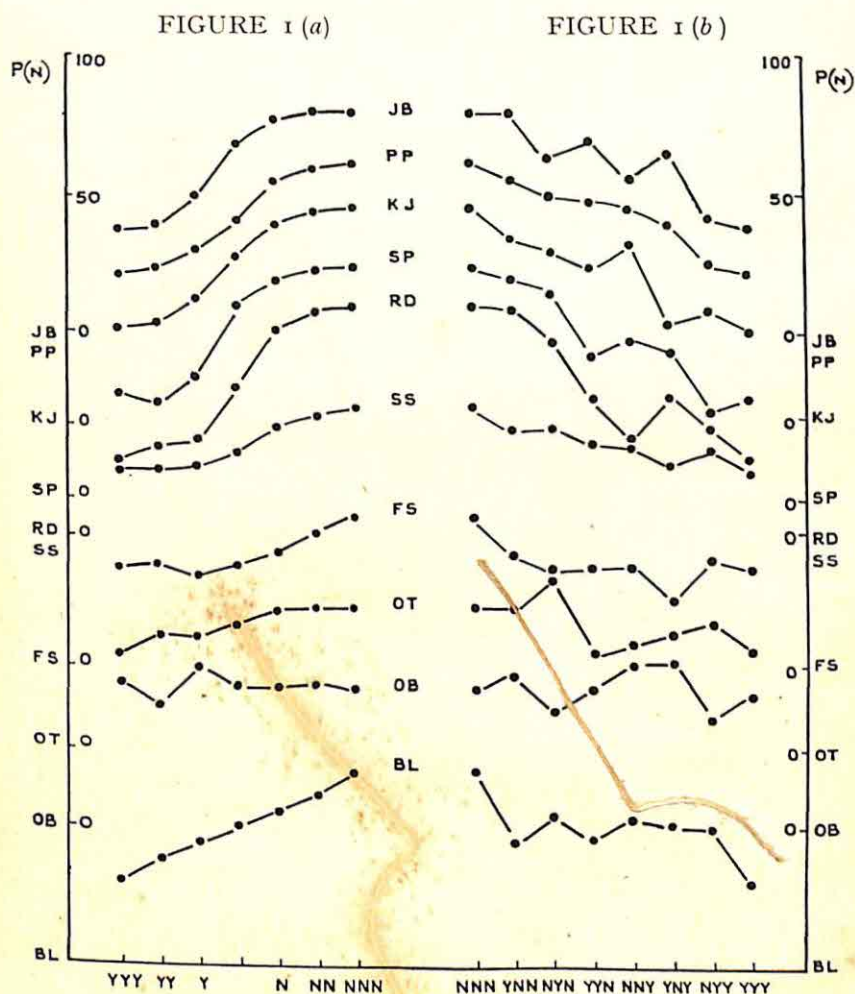
Results. The purpose of this experiment was to determine as accurately as possible the form of the non-randomness which occurs. For this purpose the method of analysis described in the introduction has been used. The results are shown in Figure 1. If the responses were random, $P(N)$ would be independent of the preceding responses and all the graphs would be horizontal straight lines. Three of the subjects, F. S., O. T. and O. B. are approximately random.

One subject, B. L., is non-random in a way suggesting very slow changes of threshold during the experiment. The other six subjects, J. B., P. P., K. J., S. P., R. D. and S. S., are all markedly non-random in a way very different from B. L. Their conditional probabilities in Figure 1*b* fall from $P(N|NNN)$ to $P(N|YYY)$ in a roughly straight line, showing that responses are dependent on the pattern of response of at least the three preceding responses; if it depended only on the last two responses we should expect

$$P(N|NNN) = P(N|YNN) > P(N|NYN) = P(N|YYN) > P(N|NNY) = \\ P(N|YNY) > P(N|NYY) = P(N|YYY),$$

giving a staircase curve in Figure 1*b* like that for subject J. B.

If $P(N)$ were dependent on the pattern of more than the three preceding responses we should expect a graph like that of subject B. L. Since $P(N|NNN)$ includes $P(N|NNNN)$ and the effects of patterns of response extending still further back, the shape of Figure 1*b* for this subject shows that he made more sustained runs of the same response than did the other subjects. The flat portion in the centre of the curve shows that $P(N)$ was nearly independent of the immediately preceding response.



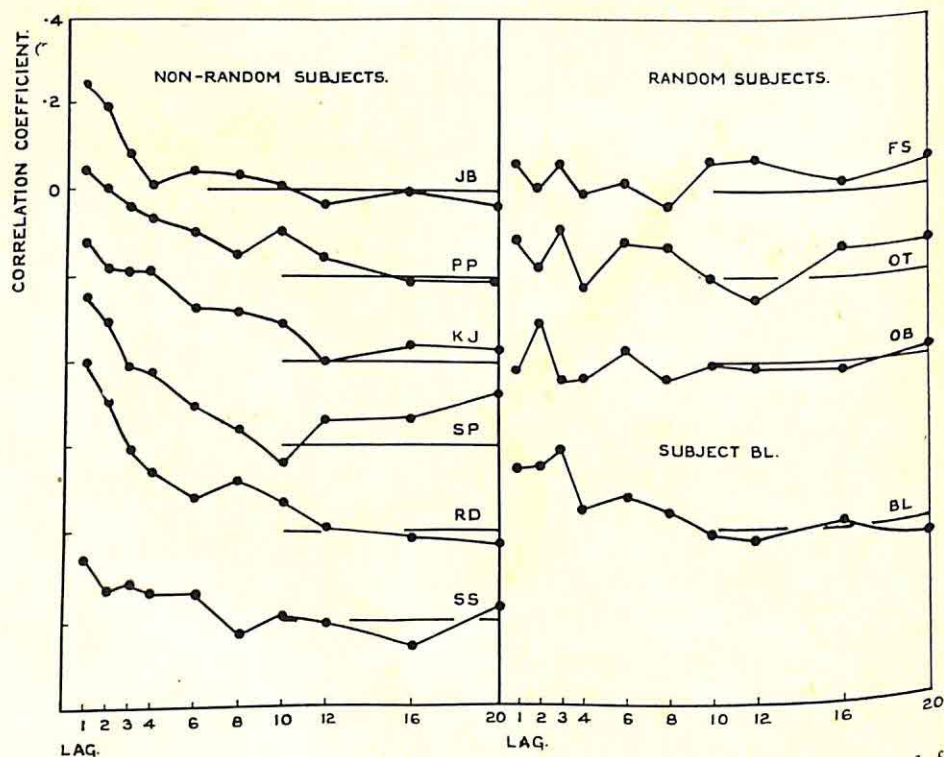
1 (a) The probability of saying "no" after runs of "yeses" and "noes" of different lengths. For the six "typical" non-random subjects the curve flattens off rapidly.

1 (b) The probability of saying "no" after the eight different patterns of three responses. In both *a* and *b* the curves for the individual subjects are staggered downwards as indicated by the zero marks on the abscissa.

Figure 1*a* confirms all these impressions. The random subjects all have flat curves. The curve for B. L. shows a continuous increase as the length of the runs increases. The other six subjects show curves flattening out after the responses three back have been fixed. Although it is easier to interpret Figure 1*a*, it is not really so informative as Figure 1*b* since the points on the curve are not independent of one another; $P(N|N)$ for example depends on $P(N|NN)$ and $P(N|NNN)$.

The results have also been analysed in terms of the correlations between responses for various intervals between them. These are shown in Figure 2. This type of analysis was used by Verplank *et al.* and by Wertheimer. Its chief advantage is that it shows up clearly any periodicity in the responses. In fact no clear periodicity can be seen either in their results or in our Figure 2. This may be because any spontaneous fluctuations were aperiodic or because the explanation of the non-randomness is the direct effect of a response on succeeding responses. Such a direct effect would in fact produce an aperiodic fluctuation of threshold when acting in conjunction with a source of random variability. This explanation is made still more plausible by a comparison of Figure 1 with Figure 2. The non-randomness seems to extend further in Figure 2 than in Figure 1. This can be explained by the "chaining" of the responses so that, if response one is correlated with response two and response two with response three, response one will be correlated with response three even if there is no direct influence between them. Chaining of this kind could only occur if there was a direct influence of one response on another. If spontaneous fluctuation were the cause of non-randomness the range of the non-randomness would be about the same in Figures 1 and 2; this is approximately the case for subject B. L. The cause of the non-randomness has, however, been investigated directly in Experiment 2.

FIGURE 2



Correlograms for the six non-random subjects, the three random subjects and for subject B. L. The individual curves are staggered downward 0.2 unit as indicated by the zero lines. Single points greater than 0.1 differ significantly from zero.

Experiment 2: The cause of the non-random sequences

Procedure. Six subjects from Experiment 1 who exhibited marked non-random behaviour were used again in Experiment 2. The initial procedure and the instructions were exactly the same. In the main part of the experiment, however, the warning bell

was sometimes rung unaccompanied by the light flash but with a masking shutter operated to prevent any auditory cues. These blank stimuli were presented in blocks of three interspersed at random throughout the experiment.

Results. The purpose of this experiment was to discover whether the cause of the non-randomness was spontaneous fluctuation of the threshold or the direct influence of one response on succeeding responses. Experiment 1 showed that a response is not appreciably affected by what happened more than three places back. If the non-randomness were due to spontaneous fluctuations of the threshold, the probability of a particular response should not be altered by forcing the subject to say "no" three times in succession by presenting a block of three blank stimuli and we should find

$$P(N) = P(N|BBB) < P(N|NNN)$$

where $P(N|BBB)$ is the probability of saying "no" to the stimulus immediately following a block of three blank stimuli. (Responses to blank stimuli are, of course, not counted in calculating any of the probabilities.) If, on the other hand, the non-randomness were due to a direct effect of one response on succeeding ones, then forcing the subject to say "no" three times to a block of blank stimuli should increase his probability of saying "no" to the stimulus immediately following the block and we should find

$$P(N) < P(N|BBB) = P(N|NNN).$$

It is essential that the blank stimuli should be presented in blocks of three so that all the three preceding responses should be fixed. If single blank stimuli had been used we should expect

$$P(N|N) > P(N|B)$$

because of the effect of the second and third preceding responses. These are more likely to have been "no" when preceding a "no" to a stimulus than when preceding "no" to a blank. It would be wrong to deduce the relative amount of spontaneous fluctuation from such an experiment as Verplank (1952) seems to have done.

The probability of saying "no" immediately before a block, $P(N||BBB)$, has been used as a check on the distribution of the blocks. If they are randomly distributed with respect to any change in the threshold, then $P(N||BBB) = P(N)$.

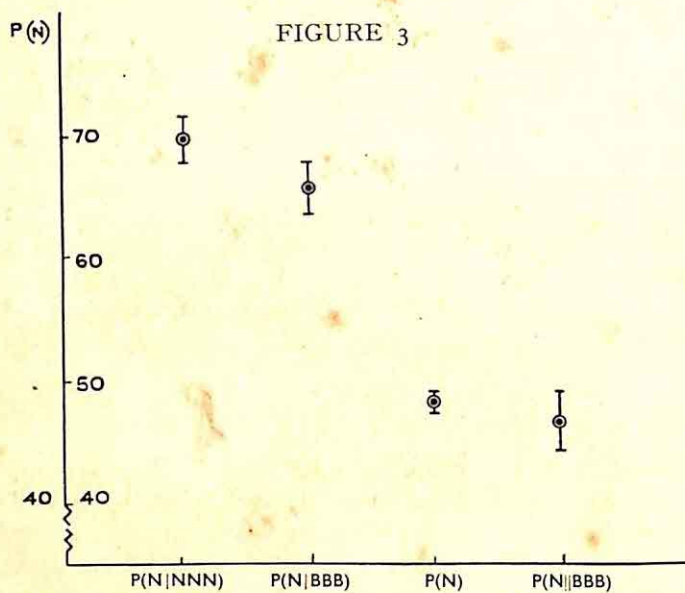
TABLE I

Subject	$P(N NNN)$	$P(N BBB)$	$P(N)$	$P(N BBB)$
S. P.	64 (67)	54 (76)	33 (552)	34 (76)
P. P.	47 (15)	44 (72)	23 (404)	32 (72)
P. P.	70 (67)	54 (52)	46 (478)	36 (52)
K. J.	83 (276)	88 (67)	74 (497)	64 (67)
J. B.	75 (44)	64 (39)	55 (275)	54 (39)
S. S.	55 (55)	67 (52)	45 (465)	48 (52)
R. D.	81 (59)	67 (42)	44 (303)	45 (42)
Weighted average $\pm$ SE.	69.9 ± 2.2	65.2 ± 2.4	48.4 ± 0.9	46.3 ± 2.6

TABLE I. The conditional probabilities of saying "no" in the various experimental situations of Experiment 2. $P(N|BBB)$ is the probability of saying "no" immediately after a block of three blank stimuli. $P(N||BBB)$ is the probability of saying "no" immediately before such a block. $P(N)$ is the probability of not seeing a stimulus in the whole experiment. $P(N|NNN)$ is the probability of not seeing a stimulus having failed to see the three preceding stimuli. Responses to blank stimuli are, of course, not included in these probabilities. The figures in brackets are the numbers of observations on which each probability is based.

The individual results are given in Table I. They have also been pooled by weighting the percentages in each row by the number of observations on which the least reliable member of the row is based. The pooled results are shown in Figure 2. No meaning can be attached to the actual probabilities in the figure although the

relationships between them are meaningful. Most of the non-randomness seems to be due to the direct influence of one response on succeeding ones and little or none to spontaneous fluctuations in the threshold. It is worth noting that any inaccuracy in the apparatus such as a drift in the intensity of the light would have had the same effect as a fluctuation of the threshold. Only two of the subjects made any responses to blank stimuli. P. P. made 9 per cent. false alarms and S. S. 2 per cent.



This figure shows the weighted average of the subjects in Table I.

Experiment 3: The effect of stimulus intensity

In Experiment 2 it has been assumed that the intensity of the stimulus is unimportant except in so far as it affects the probability of response. It has been assumed for example that a certain and an uncertain "yes" will have the same effect on the subsequent probability of response. This assumption was checked in a short experiment on four subjects. The procedure was the same as in Experiment I, but on this occasion the reaction times were measured to give an indication of the confidence of the response.

One subject gave random responses. The other three were analysed as follows. For two successive stimuli:

1. The nature of the response to the first (i.e. whether "yes" or "no") affected the average probability of response to the second by 17 per cent., but,
 2. The definiteness of the response to the first, as measured by long or short reaction times, affected the probability of response to the second by only 0.5 per cent.
 3. The average reaction time to the second when preceded by a "yes" response to the first differed from the reaction time when preceded by a "no" by 60 milliseconds in about 500; but,
 4. The average reaction time to the first when succeeded by a "yes" response to the second differs from the reaction time when succeeded by a "no" by only 2 milliseconds.
- The difference between 1 and 2 and that between 3 and 4 was only significant at the 10 per cent. level. The experiment provides no evidence for a large effect of the certainty of the response and some evidence that its effect is small.

IV

DISCUSSION

Marked non-random behaviour has been found in seven out of ten subjects. Their responses have a strong tendency to come in runs of the same response rather than alternations of different responses. In six of these subjects the effect of a response on

subsequent responses dies away rapidly having nearly disappeared by the third response back. The second experiment showed that in these six subjects the cause of this non-randomness is largely the direct influence of one response on succeeding responses rather than a spontaneous fluctuation of the threshold, which would produce runs by altering the probability of seeing in different parts of the experiment.

Wertheimer (1953) has looked for physiological rhythms in threshold data by obtaining the power functions of a considerable range of frequencies. His figures show only slight evidence of a dominant frequency, but the power function falls off rapidly for the higher frequencies. This is the type of result to be expected from our explanation of the results since too few alternations amount to a lack of higher frequencies in the fluctuation of the threshold. Correlograms are also very sensitive to fluctuations, but no sign of a dominant frequency is to be found in the work of Verplank *et al.* nor in the present experiments. One would expect a spontaneous fluctuation of threshold to be periodic. The fact that no such period in the threshold has been found makes the present explanation of the results all the more attractive.

Fluctuations can be demonstrated in continuous methods of determining thresholds (Guilford, 1927; Oldfield, 1955), but it is hard to see what relationship these have to the fluctuations of threshold which we have been investigating.

There are two other possible explanations of the non-randomness: overall trends (rather than oscillatory fluctuations) in the probability of seeing and fluctuations of the light intensity. The latter were effectively eliminated in our apparatus. They would also have shown up as a spontaneous fluctuation in Experiment 2. Any large trend would have produced results like those of subject B. L. in Figure 1 and would also have given correlations extending over considerably greater lags than are found in Figure 2. Slight trends were, however, present and were estimated by considering the difference between the probabilities of seeing in the first and last third of each experiment. The differences, d , were pooled for all subjects by taking $\Sigma(d^2/\sigma^2)$ where σ^2 is the variance of d ; this quantity is approximately distributed as a non-central χ^2 -variate with 19 degrees of freedom and non-centrality parameter $\lambda = \Sigma(\Delta^2/\sigma^2)$ where Δ is the expected value of d (see Patnaik, 1949). The value of this statistic was 91.1 from which we obtained an estimate of the weighted root mean square of 11.1 per cent. with upper and lower 5 per cent. limits of 13.4 per cent. and 7.9 per cent. Thus there is an appreciable trend in these experiments. It will, however, have little effect on the conditional probabilities. For example if the probability of seeing changed from 55 per cent. to 45 per cent. in the middle of the experiment the expected value of $P(N|N)$ would be 50.05 per cent. and of $P(N|Y)$ 49.95 per cent.

The method of analysis used in Figure 1 gives a clearer impression of the relationships involved than can be obtained by correlating responses at different intervals apart as in Figure 2 and as was done by Verplank *et al.* and by Wertheimer. For example the direct effect of one response on another might have extended only to adjacent responses. In this case it would be clearly seen from the conditional probabilities in Figure 1b since we should expect

$$P(N|NNN) = P(N|YNN) = P(N|NYN) = P(N|YYN) > P(N|NNY) = \\ P(N|YNY) = P(N|YYN) = P(N|YYY).$$

These relationships are not influenced by the chaining effect of, for example, the response three back on the response two back and of that two back on one back since the intervening responses are fixed. Correlation coefficients, however, are sensitive to this chaining effect and cannot tell one how far back the direct effect extends. The results of Verplank *et al.* are very similar and there is no reason to think that the direct effect

extended any further for their experiments on the dark adapted eye than it does in ours on the light adapted eye or that the mechanisms of the non-randomness should differ in the two cases.

Figure 1b is not independent of effects exerted from more than three places back since $P(N|NNN)$ for example is a weighted average of $P(N|NNNN)$ and $P(N|YNNN)$. Such long-distance effects have obviously caused the rather curious graph of B. L.; but they can probably be safely ignored for the other subjects, in whom the linear form of the graphs indicates that the direct effect of a previous response falls off roughly in a geometrical progression with the interval between the stimuli, since $P(N|NNN)$ and $P(N|YNN)$ are one unit apart, $P(N|NNN)$ and $P(N|NYN)$ are two units and $P(N|NNN)$ and $P(N|NNY)$ are four units apart on the graph. The conclusion that the effect of a previous response falls off sharply with its distance for the six typical non-random subjects is confirmed by the shapes of the graphs in Figure 1a. These may be compared with the almost straight line of B. L., which suggests that the effect had hardly started to fall off at the third place back and is due to long-term effects.

We have ignored the effects of the intensity of the stimulus in Experiment 2. It has been argued that the effects are due to the influence of one response on another; to justify this conclusion from Experiment 2 it must be assumed that the type of stimulus which evokes a response is irrelevant to the effect of that response on succeeding responses. This seems to be fairly justifiable under the conditions of our experiments. A short experiment using reaction time as a measure of the confidence of a response has failed to show evidence of any difference in the effect of the certainty of a particular response on the probability of succeeding responses. If stimuli of widely different intensities were used in an experiment like Experiment 2 the relationships would probably break down once the subject is able to distinguish fairly clearly between the intensities of the lights used. This is not usually the case, however, for the intensities used in the construction of frequency of seeing curves.

The instructions used were as neutral as we were able to make them and do not seem to have been a direct cause of the non-randomness. During the initial part of the experiments the intensity of light was varied as little as possible and approximately 50 per cent. frequency of seeing was obtained quite quickly. We differentiated clearly between this initial period of adjustment and the main part of the experiment; it seems unlikely that the subject's experience during the period of adjustment caused the non-randomness. In fact the experimental procedure seems less likely to have induced runs than that used in many common psychophysical procedures.

The mechanism of the non-randomness is of practical importance since the two possible mechanisms will have very different effects on the results of psychophysical experiments. For example the direct effect will tend to produce "overshooting" when thresholds are found by the method of limits, whereas spontaneous fluctuations of the threshold would merely increase the scatter of the readings. Frequency of seeing curves obtained by using grouped stimuli will have their slope increased by the direct effect, since it will increase the probabilities of seeing which are greater than 50 per cent. and reduce those less than 50 per cent.; spontaneous fluctuations would tend to flatten the curve.

When the stimuli are presented in a random order of intensities, both the direct effect and spontaneous fluctuations will decrease the slope of the frequency of seeing curve. There seems to be no method of presenting the stimuli which will eliminate the effects of either of the possible causes of non-randomness. Since random presentation of stimuli is usually used in obtaining frequency of seeing curves any estimate of the number of quanta necessary for vision obtained from such curves (e.g. Hecht,

Shlaer and Pirenne, 1942) will be an underestimate if there is any non-randomness in the responses. This is because quantum variability is necessarily random and in the case of randomised stimuli the total variability of the response is a summed function of the random and the non-random variability.

We do not wish to make too much of the possibility of non-randomness in earlier psychophysical experiments. There is no reason to suppose that the non-random variability will still be large when sophisticated subjects are used or when the subject knows that the intensity of the stimulus changes between each presentation as is usually the case with randomised stimuli. These are obviously subjects for further work. Also in the present experiments a long exposure was used deliberately, since the steep frequency of seeing curve which is characteristic of long exposures should result in a greater change in the probability of seeing for a given slow change in the threshold. It was hoped that by this means the size of the non-random effect would be increased and would be correspondingly easier to investigate. Using shorter exposures the random variability should be absolutely greater and the non-random variability proportionately less. It is impossible to know at present what will be the effect of various alterations in the experimental technique on the non-random variability and each case must be treated on its merits.

"Undershooting" in the method of limits has often been reported and implies that in this case there may be a stronger tendency to change the previous response than to repeat it.

One may tentatively consider whether this non-random behaviour is useful. It has the effect of raising the frequency of response to a more probable stimulus and of lowering it to a less probable stimulus of the same intensity. Thus the observer in effect lowers his threshold according to the probability of the stimulus without being given any information about what that probability is. This disposes of one of the difficulties encountered by those who treat sense perception as a game played against Nature in which each side tries to minimize his or her loss. Calculation of the probability of a stimulus is essential if the observer is to have any idea of his expected loss, but it is difficult to see how such a calculation could be performed in practice. The type of behaviour now reported allows the observer to behave *as if* he knew the probability of the stimulus.

ACKNOWLEDGEMENT

We are indebted to the Director General of Medical Services, R.A.F., for permission to publish this paper.

REFERENCES

- GUILFORD, J. P. (1927). "Fluctuations of attention" with weak visual stimuli. *Amer. J. Psychol.*, **38**, 534-583.
- HECHT, S., SHLAER, S., and PIRENNE, M. H. (1942). Energy, quanta and vision. *J. gen. Physiol.*, **25**, 819-840.
- OLDFIELD, R. C. (1955). Apparent fluctuations of a sensory threshold. *Quart. J. exp. Psychol.*, **7**, 101-115.
- PATNAIK, P. B. (1949). The non-central χ^2 - and F -distributions and their applications. *Biometrika*, **36**, 202-232.
- VERPLANK, W. S. (1952). Response mechanisms of the visual threshold: a methodological study. *Status report 2, Project NR 140-015*, p. 10. Office of Naval Research, U.S. Navy, 15th October, 1952.
- VERPLANK, W. S., COLLIER, G. H., and COTTON, J. W. (1952). Non-independence of successive responses in measurements of the visual threshold. *J. exp. Psychol.*, **44**, 273-282.
- WERTHEIMER, M. (1953). An investigation of the "randomness" of threshold measurements. *J. exp. Psychol.*, **45**, 294-303.

THE FUSION FREQUENCY OF FLICKER IN THE CENTRAL AND PERIPHERAL FIELD WITH PHOTOPIC LEVELS OF SURROUND LUMINANCE

BY

GEORGE ETTLINGER

*From the Psychological Laboratory, Institute of Neurology,
Queen Square, London, England*

Flicker-fusion frequencies have been measured against a photopic level of surround luminance in the four quadrants of the visual field and the critical fusion frequencies (c.f.f.s.) taken in direct vision compared with those determined at 15° and 35° in the periphery. The central c.f.f.s. of 47 subjects exceeded the peripheral values only in just over half of all the quadrants tested, although smaller differential thresholds were given almost without exception in the centre than in the periphery by the same observers and under similar conditions on a test of brightness discrimination. Likewise higher central than peripheral c.f.f.s. were recorded in the great majority of all quadrants when determinations were made against a dark surround. The contrasting types of flicker response observed against the photopic ground were shown to be relatively reproducible, and the same pattern was as a rule found in at least three quadrants of any one visual field. It was ascertained that the different patterns might result from a reciprocal shift of fusion frequencies either in central or peripheral vision.

Interpretations of the higher peripheral c.f.f.s. which adduce the effect of accentuated pupillary constriction with central stimulation were contra-indicated since such patterns of flicker perception survive after the photopupillary reflex has been abolished. Some evidence was, however, forthcoming for the view that peripheral fusion frequencies tend to exceed central values when the test object is larger than about 1.5° in angular size and more intense than the surround. It was suggested that the contrasting types of flicker response might be related to individual differences in the effect that stimulus size has on the flicker pattern.

I

INTRODUCTION

It has commonly been reported that the frequency of an intermittent light stimulus at which flicker can just be distinguished in central vision is not identical with the critical fusion frequency (c.f.f.) in the peripheral visual field. This phenomenon has, as a rule, been investigated in the dark adapted eye and with low levels of surround luminance (i.e. below 0.1 ml.). The c.f.f. for white light is under these conditions usually found to be higher in central vision than at any peripheral site if stimulus intensities exceed about 0.1 ml. at fusion; if, on the other hand, the luminance of the test object is less than about 0.1 ml. the peripheral c.f.f. is higher than the central. A similar relatively low central c.f.f. has also been reported with dark surrounds when the luminance of a 2° test object is raised above about 177 ml. at fusion (Lloyd, 1952) and for test objects exceeding 2° in angular size at levels of stimulus luminance ranging from 0.0047–21.4 ml. at fusion (Granit & Harper, 1930). Alpern & Spencer (1953) have argued convincingly that this depression of the c.f.f. with relatively intense (and perhaps also with large) test areas is to be attributed to the more pronounced response of the photopupillary reflex to central than to peripheral stimulation in the visual field under these conditions (cf. Duke-Elder, 1938, pp. 553); the smaller pupil on central testing would restrict the light incident on the retina, and the consequent reduction in the effective luminance of the stimulus would be more than sufficient to outweigh the intrinsic higher fusion threshold of this region.

The fusion frequencies of the central and peripheral areas of the visual field have been compared less often when measured against photopic and intermediate levels of surround luminance. There is, moreover, little agreement between the findings of different authors when those studies are reviewed in which the luminance of the stimulus surround has exceeded a value of 0.1 ml., in which a light : dark ratio of unity has been employed and the experimental conditions have been adequately recorded. Thus Creed & Ruch (1932) and Hecht & Verrijp (1933) reported higher central than peripheral c.f.f.s for test objects subtending 12' and 2°, and over wide ranges of stimulus and surround luminance. On the other hand the reverse relationship, namely lower rates of fusion in central vision than in parts of the peripheral field, has been ascertained by Lythgoe & Tansley (1929), Phillips (1933) and Hylkema (1942) over much of the same range of stimulus and surround luminance and for angular sizes of the intermittent light stimulus varying from 1°–5°. Riddell (1936) found a higher c.f.f. in the centre than at 10° towards the periphery for test objects of 30' and 1° angular size, but the reverse result in the same subject with larger stimuli subtending 2° and 3°. It is also noteworthy that Hecht & Verrijp alone controlled the pupillary size in these studies of the fusion frequency at levels of surround luminance greater than 0.1 ml., and that not more than 3 individuals were tested in any investigation for which the experimental conditions are fully described with the exception of the study by Lythgoe & Tansley, where the individual differences excited their comments.

In view of the contradictory outcome of previous studies, the central and peripheral flicker fusion thresholds of 47 individuals are presented. These subjects formed the control group for a clinical investigation reported elsewhere (Ettlinger, 1956). Marked variability of performance was encountered with flicker fusion against a photopic surround, and not with any other test of visual discrimination. It thus proved necessary to examine flicker-fusion thresholds under a variety of test conditions, e.g. against a darker surround.

II

METHOD

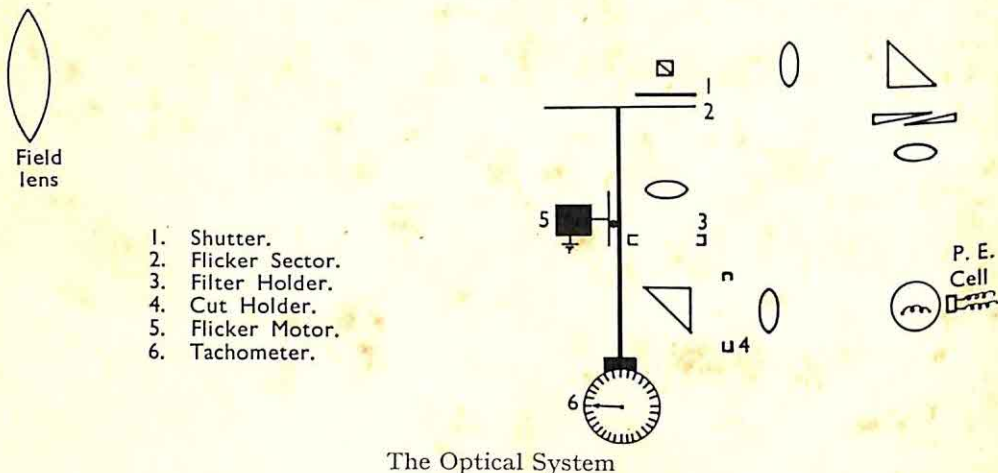
Apparatus.

An hemisphere of 1 m. radius is set vertically in a dark room. Its smooth interior surface is coated with matt-white paint and can be evenly illuminated by two 4-ft. 40-W. OSRAM "Natural" fluorescent tubes. A large corrected lens is mounted behind a circular aperture subtending 3.6° of arc at 1 m. which is let into the centre of the interior surface of the hemisphere. The intermittent light stimulus is presented from the centre of the lens. For fixation during determinations in the peripheral field of vision, small 2-V. bulbs are embedded into the plaster of the hemisphere at 15° and 35° from the centre along the diagonal meridians running from the upper left to the lower right and from the upper right to the lower left through the centre. Steady fixation is facilitated by means of an adjustable headrest and is controlled.

An optical system (Figure 1), is situated on the far side of the hemisphere from the subject. Two channels of light are taken from one source, a vertical 2-ft. 40-W. OSRAM "Natural" fluorescent tube, and after reflection by prisms are recombined in a mixing cube. The light of the background channel merely passes through neutral wedge filters; these permit the brightness of the field lens, which is seen in Maxwellian view, to be equated subjectively with that of the interior of the hemisphere. Cuts of any shape or size and filters may be inserted into their respective holders placed along the axis of the stimulus channel. The images of the cuts are seen in the centre of the field lens and in the plane of the interior surface of the hemisphere. An electrically-operated shutter, and a metal disc containing four equal sectors, cut the beam of light of the stimulus channel close to its narrowest cross-section. The distance of the sector shaft from the axis of the stimulus light beam and the latter's dimensions permit a test object of 1.5° to be fully exposed when the sector shaft has rotated through an angle of 6.6°. The four-sector disc is driven by a constant-speed motor through a "ball-and-disc" type of gearbox. A Smith's Hand Tachometer (Model A.T.H. 7) is locked to the sector shaft, and this allows a direct reading of its rate of rotation to be taken.

In order to minimize the 100 cycles component in the output of the 2-ft. 40-W. fluorescent source, the tube is supplied through a rectifier unit, and the output voltage smoothed by means of an electrolytic capacitor, the residual ripple constituting less than 8.5 per cent. of the total light output. The associated control gear of the lamp unit is retained with the addition of a variable resistor and reversing switch.

FIGURE 1



The Optical System

Subjects.

Forty-seven members of the staff of the National Hospital have been tested against a photopic surround to provide control results for a clinical investigation reported elsewhere (Ettlinger, 1956). Only two of these subjects are medical practitioners, and none has had previous experience of visual testing. The mean age of this group is 34.7 years, with a range of 17-61 years.

Procedure.

The two eyes of each subject are tested separately, and the order of testing the eyes, quadrants and positions within quadrants is systematically varied to balance practice effects. Similarly the order of the different test conditions is also varied from subject to subject in the "special" analysis. One eye is loosely covered with a large shield, but not closed. The subject is adapted to the prevailing illumination for 5 minutes prior to testing in the case of the photopic level of surround luminance, or for 15-20 minutes when testing is done in the dark.

Two ascending and two descending thresholds are obtained at each position in the visual field, using the method of minimal changes, for the phenomenal fusion of an intermittent light stimulus. The subject is always tested at first with a rate of intermittence of only about 10 c.p.s. This rate is then increased in steps of 3.3 c.p.s. until fusion is reported twice in succession, whereupon it is reduced by decrements of 0.7 c.p.s. at a time until flicker is reported, to give the first threshold. Starting from a value some 3-6 c.p.s. below this threshold, increments of 0.7 c.p.s. in the rate of intermittence are then made for a determination of "no flicker" for the second threshold. Decreasing once more from a rate some 3-6 c.p.s. above the second threshold the whole procedure is repeated. Each exposure of the test object has a duration of 1.5 seconds (timed on the electrical shutter), with 1-2 seconds intervals in between successive presentations.

Treatment of Results.

The mean value of the four threshold determinations at each site is taken as the c.f.f. for that position in the visual field. The middle 80 per cent. of the deviations of successive judgments from their mean, both in central and peripheral vision, fall in the range 0.7-3.3 c.p.s. for the group of 47 untrained observers. Since the absolute values of the fusion frequencies are of minor interest for the present investigation, the results are treated as follows:—if the c.f.f. in central vision exceeds the c.f.f. at both 15° and 35° in any

quadrant by a value greater than 0.7 c.p.s. the pattern of response for that quadrant is assigned the designation "C"; if, however, the c.f.f. in central vision is lower than that at 15° or 35° in any quadrant the response pattern for that quadrant is termed a "P" pattern.

III

RESULTS

Types of Flicker Response under Photopic Conditions of Surround Luminance for 47 Subjects.

The relationship between the c.f.f. in central and peripheral vision was determined for 47 healthy subjects. The test object subtended 1.5° of arc at the eye and was presented superimposed upon a homogeneous white field which encompassed the subject's field of vision. The luminance of the ground was 37.3 ml. and the incremental luminance of the test object was 107.3 ml. at fusion. Under these conditions the "dark" phase of the stimulus cycle was the background luminance of the field lens at 37.3 ml.

TABLE I
INCIDENCE OF 2 TYPES OF FLICKER RESPONSE IN 47 SUBJECTS

	<i>Types of Response</i>		
	<i>C</i>	<i>P</i>	<i>Total</i>
No. of eyes showing the same pattern in 4 quadrants ..	36	22	58
No. of eyes showing the same pattern in 3 quadrants ..	10	12	22
No. of eyes showing the same pattern in at least 3 quadrants	46	34	80
No. of subjects showing the same pattern in at least 3 quadrants of both eyes	17	12	29

C indicates a higher c.f.f. in central vision than at 15° and 35° in the periphery. P indicates a higher c.f.f. at 15° or 35° than in central vision.

From Table I it will be seen that a C pattern of flicker response was recorded in at least 3 quadrants of 46 eyes and a P-type of pattern in 34 eyes. The number of eyes showing the same pattern in at least 3 quadrants is, therefore, 80 out of the 94 tested. On the other hand, only 29 out of 47 subjects gave the same pattern in at least 3 quadrants of both eyes. The two eyes were never tested on the same day.

More detailed analysis of the data indicates that the C pattern of flicker perception occurs most frequently in the lower temporal, and least often in the upper nasal quadrant (χ^2 significant, $P = <0.05$). Furthermore the incidence of a higher c.f.f. at 35° than at 15° is predominant only in the lower temporal quadrant. Both of these effects are most probably to be related to the known depression of the c.f.f. in areas of the visual field near to the blind spot. On the other hand, no association between the pattern of flicker perception and either age, refractive error, sex or type of subjective appearance of the intermittent light at different sites (such as might result from differential rates of flicker adaptation, etc.) could be detected.

The ranges of the fusion frequencies taken in central vision and in the periphery of the visual field were determined separately for 34 quadrants showing C, and 38 showing P-types of flicker pattern in 18 subjects. This group of subjects comprises those individuals who were investigated for retest reliability, and only their retest

values as determined under the standard conditions described above have been analysed. C.f.f.s relating to the 15° and 35° positions in the visual field have been included individually in one frequency distribution for this analysis.

TABLE II

PERCENTAGE INCIDENCE OF DIFFERENT VALUES OF FUSION FREQUENCIES IN CENTRAL AND PERIPHERAL VISION FOR THE CONTRASTING RESPONSE PATTERNS OF 18 SUBJECTS

	<i>Types of response</i>	<i>Below 40 c.p.s.</i>	<i>40-49 c.p.s.</i>	<i>50-59 c.p.s.</i>	<i>60-69 c.p.s.</i>	<i>Above 69 c.p.s.</i>
Central	C	0	53.8	15.4	23.1	7.7
Vision	P	0	50.0	41.7	8.3	0
15° and 35° in periphery . .	C	31.4	32.9	32.9	2.8	0
	P	0	18.7	57.3	18.7	5.3

Although there is considerable overlap between the ranges associated with C and P patterns of response both in central and peripheral vision, it is evident from Table II that central c.f.f.s are not only more frequently higher, but peripheral c.f.f.s are also lower in quadrants showing C patterns as against P types of response. Thus 30.8 per cent. of c.f.f.s taken in central vision exceed 60 c.p.s. when related to the C response patterns as against 8.3 per cent. when associated with P patterns. Similarly the proportion of fusion frequencies determined in the peripheral field of vision that exceeds 60 c.p.s. in quadrants showing the C class of pattern is 2.8 per cent. compared with 24.0 per cent. in quadrants showing the contrasting response. Again 64.3 per cent. of peripheral c.f.f.s related to C patterns, compared with 18.7 per cent. of P-types of response, fall below 50 c.p.s. C-types of pattern may then differ from the P class in that either the central c.f.f.s reach higher values or the peripheral fusion frequencies are depressed.

The main finding, of clear-cut differences in the relationship of the fusion frequency to the eccentricity of stimulation in the visual field both for different subjects and not infrequently for the two eyes of a single individual, should be contrasted with the almost uniform response patterns of the same subjects in brightness discrimination. For employing the same apparatus and test conditions the differential limen for brightness was smaller in central vision than at 15° in 374 out of 376 quadrants, and greater at 15° than at 35° in only 3 out of 376 quadrants. A similarly consistent relationship between discrimination and eccentricity of stimulation was found on tests of minimum separable acuity and tachistoscopic acuity.

Test-Retest Reliability of the Flicker Patterns.

Retaining the same experimental conditions and procedure, 18 subjects were retested on one eye after intervals of time ranging from 1-12 months. Eight of these subjects were selected by means of a random procedure, the remainder because their fusion frequencies conformed to particular classes of flicker response. The order of testing quadrants and positions within each quadrant was reversed for 7 subjects.

The pattern of flicker response was found to be changed in 15 out of 72 quadrants. Nevertheless statistical analysis gives evidence of a significant relationship between the test and retest results ($\chi^2 = 22.61$, 1 d.f., $P = < 0.01$). The reversal in the order of testing did not give rise to selectively lower test-retest reliability. Comparing the

response patterns for all 18 subjects the C pattern of flicker perception is more stable than the P pattern. For in only 14.8 per cent. of the 27 quadrants originally showing a C pattern was a P-type of response found on retesting, whereas a reversal from P to C patterns occurred in 24.4 per cent. of quadrants. No such differences are in evidence for the group of 8 unselected subjects which, however, contains only a small sample (8) of C patterns.

Types of Flicker Response under Scotopic Conditions of Surround Luminance.

Thirteen subjects were retested on the same apparatus, but against a surround luminance of 0.00118 ml. and after a period of 15-20 minutes dark adaptation. Two conditions of stimulation were investigated. In the first, the intermittent light was of the same angular size (1.5°) and luminance (107.3 ml. at fusion) as had been employed at the photopic level of surround luminance. In the second, a test object of $45'$ angular size and 1.073 ml. luminance at fusion was viewed against the dark ground. The luminance of the dark phase of the stimulus cycle was approximately one-tenth of the surround luminance under both test conditions.

TABLE III
EFFECT OF DARK SURROUND ON FLICKER PATTERNS OF
13 SUBJECTS

Test conditions: Scotopic I—test object of 1.5° angular size and 107.3 ml. luminance at fusion. Scotopic II—test object of $45'$ angular size and 1.073 ml. luminance at fusion.

	Standard response	Scotopic I		Scotopic II	
		C	P	C	P
C	25	25	0	23	2
P	27	23	4	21	6

The findings are tabulated in Table III. C patterns of response had been recorded under photopic conditions of surround luminance in 25 quadrants, and these patterns are reproduced almost perfectly against the dark ground with both sizes and luminances of the test object. Moreover a clear majority of the 27 quadrants showing P classes of response against the photopic surround were found to show the C pattern under scotopic conditions of the surround. Only in 4 quadrants (of 2 subjects) were P patterns still recorded with the more intense test object, and in 6 quadrants (of 4 subjects) with the less intense stimulus. The difference between the proportion of C responses under the standard and both scotopic conditions of surround luminance can be demonstrated statistically to be significant beyond the 5 per cent. level of confidence.

Analysis of the Flicker Response Patterns Occurring under Photopic Conditions of Surround Luminance.

The same 13 subjects were required to give flicker fusion judgments against the standard surround luminance of 37.3 ml. and with intermittent stimuli of 107.3 ml. incremental luminance at fusion but for different sizes of test object. Under these conditions the "dark" phase of the stimulus cycle is equal to the background in luminance.

When test objects of $45'$ and 3.6° angular size were presented the C pattern of response was observed to remain relatively stable with both increase and decrease in the size of the test object from the 1.5° standard size (8 per cent. pattern reversal in each case). Those quadrants, on the other hand, in which a P pattern of response had originally been recorded, showed about the same degree of pattern reversal (25.9 per cent.) when tested with a 3.6° object as on retesting under the standard conditions, but a more marked shift (48.1 per cent.) towards the C pattern in the case of the $45'$ object.

Since no previous investigator has used a "dark" phase in the stimulus cycle equated for luminance with the photopic background, further testing was carried out with the luminance of the dark phase having a value approximately 5 per cent. of that of the surround. Reverting to the standard size of the test object, the 13 subjects were again tested but with the luminance of the dark phase of the stimulus cycle and also of the immediate surround of the test object (i.e. of the field lens) reduced to 1.9 ml. A circular object of 1.5° angular size was therefore seen as bright (at 107.3 ml. at fusion) against an immediate surround or annulus of 1.05° width. This annulus, having a luminance of 1.9 ml., was itself surrounded by the hemisphere of 37.3 ml. luminance.

The results indicate that the instability of both the C and P patterns (20 per cent. and 44.4 per cent. shift) is more marked than on retesting under standard conditions. This increased tendency to pattern reversal may be related to the effect of the darker immediate surround or to the change in the dark phase. These two variables could not be isolated on the apparatus using a 1.5° test object. However, on retesting in 8 quadrants which had originally shown P patterns, with a 3.6° test object and a dark phase of 1.9 ml. but without the dark annulus, a deviation from the P pattern was not recorded in a single instance.

In view of the findings of Alpern & Spencer (1953) it was considered advisable to control the pupillary response in a further investigation. Fusion frequencies were determined for 6 subjects under the standard conditions of surround and stimulus luminance and with test objects subtending $45'$, 1.5° and 3.6° of arc. Three of the subjects (showing 8 C and 4 P patterns originally with a 1.5° test object) were tested 2 hours after their pupils had been constricted with eserine, and 3 subjects (showing 4 C and 8 P patterns originally) were given homatropine 2 hours before testing. Results are therefore available for 24 quadrants in all.

TABLE IV
EFFECT OF CONSTANT PUPILLARY SIZE ON THE FLICKER PATTERNS
OF 6 SUBJECTS FOR 3 SIZES OF TEST OBJECT

	Standard response	$45'$		1.5°		3.6°	
		C	P	C	P	C	P
C	12	12	0	11	1	8	4
P	12	6	6	4	8	0	12

From the data of Table IV, the two contrasting patterns of flicker response may be seen to survive abolition of the photo-pupillary reflex. A slight shift towards the C-type of pattern can, however, be detected with the 1.5° test object. For not only is the incidence of reversal from P to C patterns (33.3 per cent.) higher than on

retesting under standard conditions (24.4 per cent.) but the frequency of change from C to P patterns is at the same time reduced (8.3 per cent. against 14.8 per cent.). This trend in favour of the C-type of flicker perception is accentuated with the 45' test object. Thus half the quadrants originally showing P patterns have changed to the other pattern, whereas no deviation away from the C-type of response is recorded. A shift in the converse direction occurs with the test object of 3.6° angular size. For 33.3 per cent. of C patterns are changed but all the P-types of responses are reproduced. No association between any one kind of pattern reversal and specifically either miosis or mydriasis was observed.

IV

DISCUSSION

Contrasting relationships between the flicker fusion frequency and eccentricity of stimulation can be observed in a large group of individuals over a range of test objects extending from 45'-3.6° in angular size when a photopic level of surround luminance is retained. However, as anticipated from previous work, differences in the pattern of flicker perception are largely if not entirely resolved when testing is conducted against a dark surround (0.00118 ml.). Both of the contrasting photopic response patterns, and more especially the P-type of flicker perception, appear to lack real stability. Thus a pattern reversal occurs in about 21 per cent. of quadrants when the test and retest conditions are identical. Nevertheless the two-tail probability that the retest results conform to the hypothesis that C and P patterns are equally likely to be recorded on retesting, is only 0.0004 in the case of quadrants originally showing C, and 0.001 for quadrants originally showing P patterns of flicker perception. An explanation must therefore be sought for the occurrence of two relatively reproducible kinds of flicker response under constant photopic experimental conditions in subjects who almost uniformly give increased thresholds with greater eccentricity of testing on brightness discrimination and other tests.

It is unlikely in view of the similar anomalies of flicker perception reported in studies employing light from filament sources that the present differences of patterns are the direct result of the characteristic wavelength distribution of the light emitted by the fluorescent tubes. Similarly the 100 c.p.s. alternation of the surround illumination is most probably of no immediate consequence for the origin of these patterns. Not only is this frequency well above the maximal flicker fusion frequency observed, but it may also be doubled to 200 c.p.s. (by suitable arrangements of capacitors) without giving rise to changes in the contrasting flicker patterns of 2 subjects. Finally, particular attention was directed to the possibility of interference with flicker perception from "beats" between the residual source and intermittence frequencies, but no evidence whatsoever for such interference could be detected. None of the test conditions specific to the present investigation appear to account therefore for the occurrence of the contrasting flicker patterns.

The interpretation advanced by Alpern & Spencer (1953) to explain the depressed central fusion frequencies with relatively intense and perhaps also large test objects against a dark surround, makes reference to the more marked photopupillary constriction with central stimulation. Differential activation of this mechanism can not, however, be adduced to account for the origin of the two distinct patterns under the standard photopic test conditions. For higher c.f.f.s in the periphery than in direct vision may still be recorded after the photopupillary reflex has been abolished, and in fact are observed relatively more frequently in the case of large (3.6°) test objects when the pupillary size remains constant than when the photopupillary reflex is intact.

This study gives some evidence of the greater prevalence of C types of response with small ($45'$) test objects and of P patterns with large (3.6°) areas of intermittence—although the pattern shift is by no means complete with either size of stimulus. This finding accords in general with the results of Granit & Harper (1930) for dark surrounds, and of Riddell (1936). The exact value of the background luminance employed by Riddell in obtaining the relevant results is not given, but may be calculated to be a little above 0.1 ml. It is also noteworthy that Creed and Ruch (1932), who reported higher c.f.f.s in central than peripheral vision, employed on intermittent light subtending only $12'$, whereas Phillips (1933) and Hylkema (1942) found the reverse pattern of flicker perception on presenting areas of 1.5° or greater.

The results of Lythgoe & Tansley (1929), namely higher peripheral c.f.f.s with a 1° square test object, and of Hecht & Verrijp (1933), namely the reverse pattern with a 2° area, do not conform to the principle that areas of intermittence greater than about 1.5° are associated with maximal fusion frequencies in the peripheral field. Since in both studies the luminance of the test object at fusion was equated with that of the photopic surround (whereas the other authors referred to have presented stimuli of greater intensity than the background) stimulus size may have a different effect on the flicker pattern when the surround and test object are of equal luminance. This interpretation is supported by the reversal shown by 2 out of the 3 subjects investigated by Lythgoe & Tansley from P to C patterns of responses when tested with an object of the same size and luminance but against a surround luminance reduced from equality with the test object at 8.4 ml. to 0.93 ml.

At any rate it appears evident from the findings that the same pattern of flicker response is not shown consistently by all individuals nor in both the eyes of a single individual under certain test conditions which involve a photopic level of surround luminance. Individual differences were also marked in the investigation of Lythgoe & Tansley under photopic conditions although these authors do not report instability of the response in time. It may well be that uniform patterns of flicker perception could be recorded with sizes of test object larger than 3.6° or smaller than $45'$. The contrasting response patterns with a 1.5° angular size of test area could then be related to individual differences in the transition point between C patterns with small, and P patterns with large test objects. The basis for such differences (already observed by Granit & Harper, 1930, against a scotopic level of surround luminance) is not known but may, perhaps, be attributed to individual variations in the relative distribution of the rods and cones. Confirmation for this hypothesis could be obtained from a comparison of flicker pattern and extent of perimetric field for coloured objects. The present findings likewise do not contribute towards a further explanation for the emergence of higher c.f.f.s in the peripheral than central field with large test objects against a photopic surround.

ACKNOWLEDGEMENTS

The late Dr. L. C. Thomson of the Institute of Ophthalmology kindly gave me help in designing the optical system and made available certain optical equipment on loan. I also wish to thank Mr. H. B. Morton, who helped with the design and construction of much of the electrical apparatus; Dr. W. A. Cobb, Mr. R. L. Gregory and Dr. K. H. Priam for their assistance in preparing this paper; and Dr. Eliot Slater for special facilities. My especial gratitude is due to Prof. O. L. Zangwill for his constant encouragement and advice.

REFERENCES

- ALPERN, M., and SPENCER, R. W. (1953). Variation of critical fusion frequency in the nasal visual field. *Arch. Ophthalm.* N.Y., 50, 50.

- CREED, R. S., and RUCH, T. C. (1932). Regional variations in sensitivity to flicker. *J. Physiol.*, **74**, 407.
- DUKE-ELDER, Sir W. S. (1938). *Text-book of Ophthalmology*, **1**. London.
- ETTLINGER, G. (1956). Sensory deficits in visual agnosia. *J. Neurol. Psychiat.* In the press.
- GRANIT, R., and HARPER, P. (1930). Comparative studies on the peripheral and central retina: II. Synaptic reactions in the eye. *Amer. J. Physiol.*, **95**, 211.
- HECHT, S., and VERRIJP, C. D. (1933). Intermittent stimulation by light: III. The relation between intensity and critical fusion frequency for different retinal locations. *J. Gen. Physiol.*, **17**, 251.
- HYLKEMA, B. S. (1942). Fusion frequency with intermittent light under various circumstances. *Acta. Ophthal., Kbh.*, **20**, 159.
- LLOYD, V. V. (1952). A comparison of critical fusion frequencies for different areas in the fovea and periphery. *Amer. J. Psychol.*, **65**, 346.
- LYTHGOE, R. J., and TANSLEY, K. (1929). The adaptation of the eye: Its relation to the critical frequency of flicker. *M.R.C. Spec. Rep. Ser.*, No. 134.
- PHILLIPS, G. (1933). Perception of flicker in lesions of the visual pathways. *Brain*, **56**, 464.
- RIDDELL, L. A. (1936). The use of the flicker phenomenon in the investigation of the field of vision. *Brit. J. Ophthal.*, **20**, 385.

MISCELLANEA

THE BAR-PRESSING RECORDER

BY

J. R. TROTTER

*From the Institute of Experimental Psychology,
University of Oxford*

The simple weighted bar, as commonly used in bar-pressing experiments with rats, is shown schematically in Figure 1a, together with its force-displacement graph. The possible modes of pressing the knob, and the reactions of the bar are described in a previous paper (Trotter, 1956). The simplest and probably the most common way of recording the rat's behaviour towards the knob is to count the number of times a maximal displacement has occurred. The usual name for this variable is "number of responses." Data on the intervals between "responses" may be obtained by recording each "response" on a kymograph. Skinner (1938) used a cumulative trace whose slope indicated "frequency of responses." These systems deal only with maximal displacements. If a stylus is actuated by the knob a trace may be made showing the magnitude and the duration of any displacement, whether maximal or smaller. The last system provides all the data which can possibly be obtained from the simple weighted bar.

The deficiency of the simple weighted bar, no matter what recording system is used, is that it provides meagre data which, owing to an ambiguity of the term "response," have sometimes been misinterpreted. In the previous paper a new device, the force recorder, is described. This provides a temporal record of force applied to the knob. Using the terminology of the previous paper, the force recorder gives a record of output, and the simple weighted bar provides a selection of data concerning the result.

The force-displacement graph in Figure 1b shows that the equation of motion of the knob (assuming a frictionless bar of negligible mass) is $x = Cf$. This equation applies under any conditions of position and downward force on the knob. A press on the knob consists of a single phase, and is a mechanically simple event. The graph in Figure 1a has a discontinuity at $f = W$, indicating three equations of motion:—

$$(1) (f < W) \text{ steady state} \dots x = 0.$$

$$(2) (0 < x < s) \dots x = \frac{g}{W} \iint (f - W) dt. dt.$$

$$(3) (f > W) \text{ steady state} \dots x = s.$$

A press may thus consist of five distinct phases:—Initial phase, lifting phase, holding phase, lowering phase and final phase. There are two reasons why the force recorder is *a priori* preferable to the simple weighted bar:—A press on the former is a mechanically, and physiologically simpler event than a press on the latter; the former gives more complete and more useful data on the record.

However, during the last 20 years hundreds of experiments have been performed using the simple weighted bar. Experiments using the force recorder are not directly comparable with these because of the difference in mechanical and physiological events. For the sake of continuity of research a combination of the two instruments is required, a device which simulates the simple weighted bar, yet provides the full data given by the force recorder.

Figure 1c shows the principle of the bar-pressing recorder used by the writer. Its three equations of motion, indicated by graph 1c, are:—

$$(1) (f < W) \text{ steady state} \dots x = Cf.$$

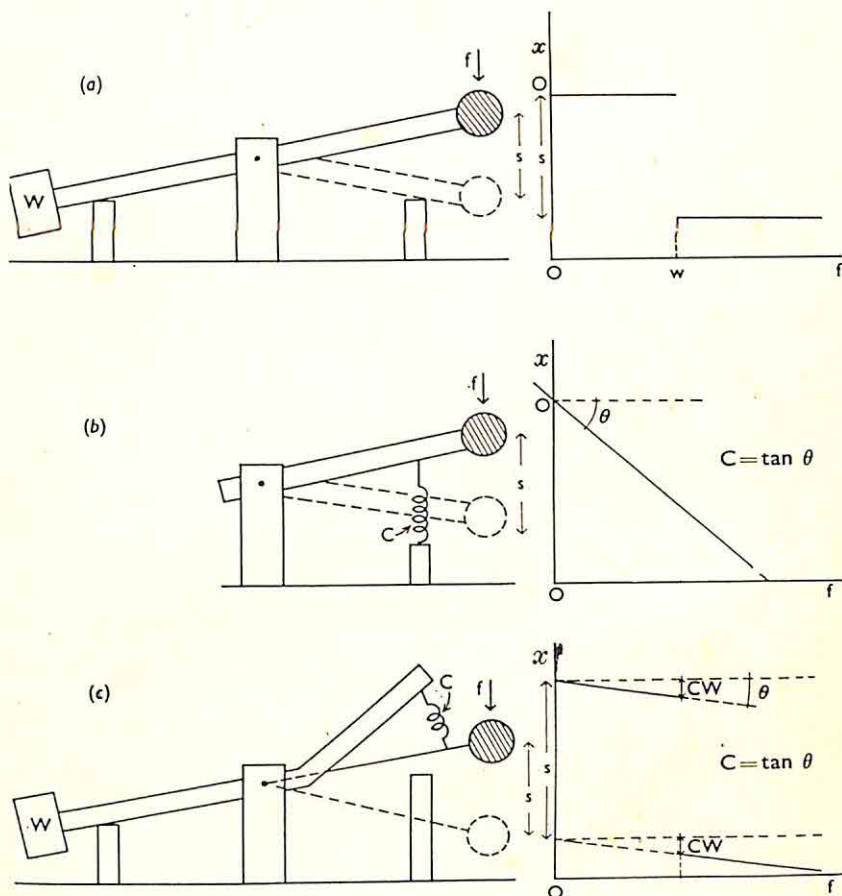
$$(2) (CW < x < s + CW) \dots x = Cf + \frac{g}{W} \iint (f - W) dt. dt.$$

$$(3) (f > W) \text{ steady state} \dots x = s + Cf.$$

Since $C \approx 1 \text{ mm./100 gm.}$, and $s \approx 1 \text{ cm.}$, Cf is negligibly small compared with s , for any force used by a rat, and the force-displacement graph approximates to that of the simple

weighted bar. By the use of a lightweight capillary pen the resonant frequency of the knob is raised to about 40 c/s, so the knob can follow the quickest variations of force which can be expected. The knob is rigid laterally. Movement of the bar is made almost noiseless by means of padded stops, at the expense of slightly rounding the corners of the force-displacement graph.

FIGURE 1



Some bar-pressing devices and their force-displacement graphs

(a) Simple weighted bar. (b) The force recorder. (c) The bar-pressing recorder, a compound of (a) and (b).

The bar is centrally pivoted and the angle of rotation is presumed negligibly small. W = weight, C = compliance, s = maximal displacement, x = displacement, f = force applied vertically downwards.

The specification of the bar-pressing recorder is as follows:—

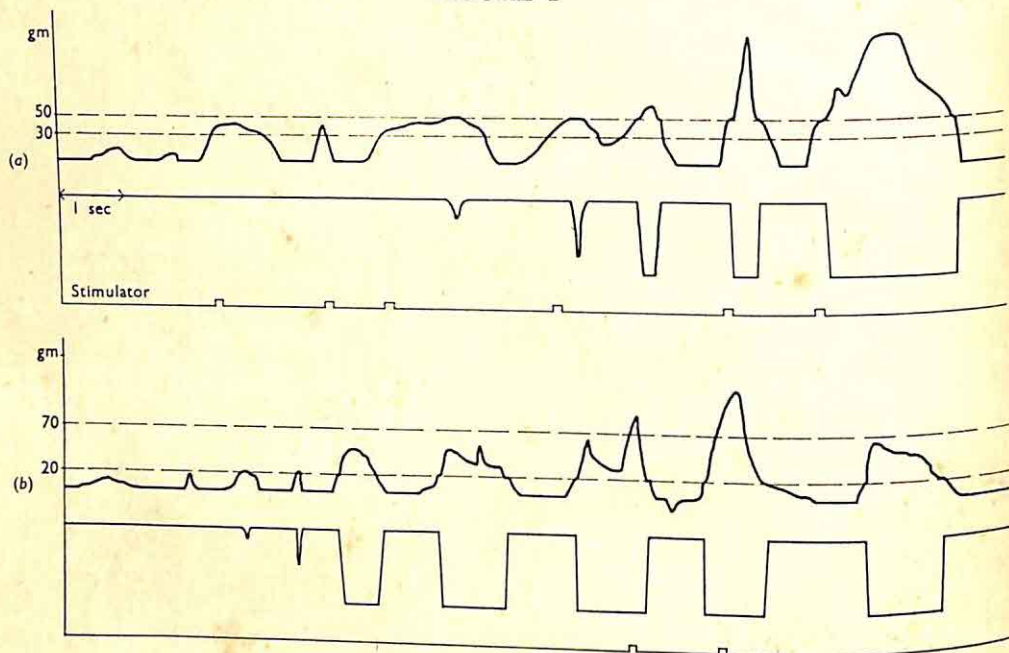
Situation. The experimenter may fix the force required to move the bar at any multiple of 5 gm. up to over 100 gm. by means of a set of weights. Any movement of the bar breaks number 1 contact, and a maximal displacement makes number 2 contact. Independently of movement of the bar, any applied force exceeding 2–3 gm. (nominally zero) breaks number 3 contact, and at any force chosen by the experimenter by setting a pointer on a scale number 4 contact is made. Any of the contacts may be used to operate, stop or reset a rewarding device or a sensory stimulator.

Output. A temporal record of vertically applied force, up or down (with negligible disturbance by movement of the bar) on a scale of about 1 cm./50 gm.

Result. A temporal record of displacement of the bar, and of delivery of reward or sensory stimulus.

Sample records are shown in Figure 2. Records are measured by using a transparent graticule with appropriate vertical and horizontal divisions.

FIGURE 2



(a) $W = 50$ gm. (b) $W = 20$ gm.

No. 1 contact resets the rewarding device. ($x = 0$).

No. 2 contact operates the rewarding device. ($x = s$).

No. 3 contact resets the stimulator. ($f = 0$).

No. 4 contact operates the stimulator at:—

(a) $f = 30$ gm. (b) $f = 70$ gm.

In brief, the bar-pressing recorder simulates (to the rat) a simple weighted bar, but provides data on the physical properties of behaviour which the latter can not, and which have been neglected by investigators.

REFERENCES

- SKINNER, B. F. (1938). *The Behavior of Organisms*. New York.
- TROTTER, J. R. (1956). The physical properties of bar-pressing behaviour and the problem of reactive inhibition. *Quart. J. exp. Psychol.*, **8**, 97-106.

A Note on the Paper* by Richard A. Littman and Edward A. Wade entitled "A Negative Test of the Drive-Reduction Hypothesis"

BY

J. A. DEUTSCH

From the Institute of Experimental Psychology, University of Oxford

Littman and Wade's conclusions in their article appearing in Vol. 7 (pp. 56-66) of this journal should be viewed in the light of the fact that Coppock (1950, 1951) reports a very similar experiment with a contrary result. A comparison of the procedural differences in these experiments is valuable as they may have an effect on some of Littman and Wade's conclusions.

Coppock used 80 rats. These were clamped into a snugly-fitting crate with head and neck protruding. Their tail was then shocked by the application of alternating current once per minute for 5 seconds. The illumination was blinked for 20 rats for 1 second, $\frac{1}{2}$ a second before the end of the shock, for another 20 rats, at the termination of the shock and for another 20 rats, 2 seconds after the end of the shock. The remaining 20 animals were given no blink. All the animals were given 30 shocks on each of three successive days.

During three 30-minute test periods a photo-electric relay blinked the light as long as the head position of the animal in the training apparatus was more than 22 degrees off centre. Half the animals in each group were reinforced when they turned to the preferred side. The other half was rewarded on the non-preferred side.

Significant group differences were found only for animals reinforced only on the non-preferred side. There was a greater total duration of holding the head in the "reward" position in the group for which the light had coincided with the end of the shock ($P = 0.02$ when this group is compared with that for which the blink had preceded the end of the shock and $P = 0.01$ when it is compared with the group where there was a 2-second gap between the termination of the shock and the blink). This result tends to substantiate the prediction from Hull's theory as stated by Littman and Wade.

In Littman and Wade's experiment the animals were placed in a completely dark tube (or so we are left to infer). The animals in the experimental group ($N = 7$) were shocked through the tail by the administration of alternating current for half a second. A tenth of a second later, a light was shone at the rat for one second.

In the test situation, the animal was placed in an illuminated Skinner box. Above, the lever in this box was the milk-glass screen through which the light has been directed into the animal's eyes in the training sessions. The animals did not differ significantly from their controls in failing to acquire the lever-pressing habit which would have produced the light for them.

There are two differences in these two experiments deserving attention. The first involves the use of light in an experiment involving electric shock, and the second the removal of the animal from the original situation to test the effectiveness elsewhere of the secondary reward, if any. It is to be noted that Coppock blinked a light in his experiment; Littman and Wade switched it on. There is the general point that Littman and Wade's rats may have been mildly punished by this procedure. The mild punishing effect of this light may have swamped any secondary-reward value which it may have acquired. Certainly a test of the assumed initial neutrality of the stimulus should have been made. There is the more particular point which involves the use of a light as a stimulus in an experiment involving electrical stimulation. The retina is extremely sensitive to electrical stimulation and the production of electrical phosphenes during the shock would mimic just the kind of diffuse visual stimulation used at the termination of the shock by Littman and Wade. This presents a possible source of experimental error and one which Coppock's design to some extent avoids.

The second point concerns what has been called "trans-situational" reinforcement. Coppock tested his animals in the same apparatus as he trained them. Littman and Wade shift their animals along with one aspect of the training situation. We do not, of course,

*(*Quart. J. exp. Psychol.*, 7, 56-66, 1955)

know how far a rat would be able to identify this same light under the conditions of general illumination. However, Littman and Wade, if they can show that the degree of resemblance was sufficient, may claim that they have shown in conjunction with Coppock's study that any secondary reinforcer acquired in this situation is not "trans-situational". This may very well be the case where fear has been the motivation, but their experiment has not even demonstrated this.

To draw this conclusion we should first have to establish that the light has secondary reinforcing properties in the training situation itself and as has been argued above, there is some doubt about this. However, a test of this possibility would make an interesting experiment.

REFERENCES

- COPPOCK, H. W. (1950). An investigation of secondary reinforcing effect of a visual stimulus as a function of its temporal relation to shock termination. Unpublished doctoral dissertation, Indiana University, 1950.
- COPPOCK, H. W. (1951). Secondary reinforcing effect of a stimulus repeatedly presented after electric shock. *Amer. Psychol.*, 6, 277.
- LITTMAN, R. A., and WADE, E. A. (1955). A negative test of the drive reduction hypothesis. *Quart. J. exp. Psychol.*, 7, 56-66.

Reply to the Note by J. A. Deutsch

BY

RICHARD A. LITTMAN and EDWARD A. WADE

From the University of Oregon and Tufts University

Deutsch has made four criticisms of our paper, one is based on the work of Coppock; another deals with an alleged interaction of electrical and photic stimulation; a third suggests that the light stimulus has aversive properties which we did not satisfactorily control; and a fourth suggests that the mode of testing was inappropriate. We shall consider these in reverse order.

Mode of Testing. It is still the case that Hull's formulation permits a "trans-situational" test condition. This matter was considered in the original report of our investigation, and Deutsch has not shown the reasoning to have been faulty. Instead, he has objected to the procedure because we do not guarantee that the training and test situations resemble one another. It is startling, to say the least, to see a defence of Hull couched in terms of "resemblance" or an "identification process." Hull's explicit formulation of stimulation is physicalistic and is designed to permit the deduction, as theorems, in his system, of precisely such events and outcomes as identifying and resembling. Beyond using physicalistic criteria like brightness, geometric shape, etc., we know of no way to "show that degree of resemblance was sufficient." If Deutsch's suggestions were to be followed out, we fear that only results favourable to Hull's position would afford good evidence that a resemblance was noted.

Phosphenes and Shock. This criticism is two-pronged: Motokawa's data certainly shows that electronical phosphenes are generated during shock. It should be noted, however, that in all his work, Motokawa placed the electrodes on the head. Whether a shock to the tail would generate phosphenes is simply not known; it, too, would make an interesting experiment. But, it is hardly fair to criticize a procedure on the possibility that something about which no one knows anything might be true. More important, however, is the fact that Coppock is in the same predicament. According to Deutsch, phosphenes-shock interaction is something which his "design (only) to some extent avoids." If, then, Coppock is subject to the same errors, his data cannot provide a legitimate basis for criticism.

Light as Aversive Stimulus. It is pointed out by Deutsch that the light stimulus we used shone steadily for the full interval; in contrast, Coppock used a blinking light. The implication is that a blinking light is less "punishing" (aversive) than a continuous light. It is difficult to know just what the nature of the punishment from the light is supposed to be. There is no reason to believe that a steady light is more distressing than an intermittent one; indeed, for equal brightness, flicker may likely be more distressing. What is

more likely as a factor is the brightness of the light; while Coppock is not clear as to what his "shielded light" is, ours was probably brighter (though of the same wattage) since it was placed within a reflecting cone which doubtless served a condensing function.

However, while we did not have a pre-training control of the neutrality of the light (neither did Coppock) Group V serves the same function. If the light were aversive, then during the test situation this group's operant level should have been depressed. Had it been depressed Deutsch's point would be more cogent; as it turned out, Group V actually ran a slightly higher rate though not significantly so.

Coppock's Research. We shall assume that neither Coppock's work nor ours is seriously endangered by the last two criticisms. We did not discuss Coppock's results in our paper, because they are ambiguous at best and, one would think, negative at worst. Here is why: Hull speaks of an interval of "20 seconds or so." For Coppock all six conditions, preferred or non-preferred by preceding, simultaneous, or following, were within this range. Assuming that the "following" conditions should, at the very least, yield significant results, we have only one out of two actually turning out that way. This becomes only two out of four in favour of the hypothesis, if we consider why the non-preferred rather than the preferred showed an increase in relative frequency. The interpretation of the results is therefore most difficult. Coppock's (1954) subsequent research is in the same negative direction as far as drive-reduction theory is concerned. If Coppock's results and ours are totalled up, the preponderance for all conditions is against the hypothesis.

REFERENCE

- COPPOCK, H. W. (1954). Stimulus preceding electric shock can acquire positive reinforcing properties. *J. comp. physiol. Psychol.*, **47**, 109-113.

INDEX OF AUTHORS

ANDJUS, R. K., with Knöpfelmacher, F., Russell, R. W., and Smith, A. U.	15
BEACH, F. A., with Jordan, L.	121
BELBIN, E., with Cane, V. R.	45
BROADBENT, D. E.	145
BROWN, J.	134
BRUNER, J. S., with Jeeves, M. A.	107
BULMER, M. G., with Howarth, C. I.	163
CANE, V. R., with Belbin, E.	45
CHEESMAN, G. H., with Townsend, M. J.	8
DAVIS, R.	24
DAVIS, R.	39
DEUTSCH, J. A.	185
ETTLINGER, G.	172
FRAISSE, P., with Vautrey, P.	114
HICK, W. E., with Welford, A. T.	39
HOWARTH, C. I., with Bulmer, M. G.	163
JEEVES, M. A., with Bruner, J. S.	107
JORDAN, L., with Beach, F. A.	121
KAY, H., with Skemp, R.	153
KHAIRY, M., with Knöpfelmacher, F., Russell, R. W., and Yudkin, J.	54
KNÖPFELMACHER, F., with Andjus, R. K., Russell, R. W., and Smith, A. U.	15
KNÖPFELMACHER, F., with Khairy, M., Russell, R. W., and Yudkin, J.	54
LITTMAN, R. A., with Wade, E. A.	186
OLDFIELD, R. C.	44
PROVINS, K. A.	79
RUSSELL, R. W., with Andjus, R. K., Knöpfelmacher, F., and Smith, A. U.	15
RUSSELL, R. W., with Knöpfelmacher, F., Khairy, M., and Yudkin, J.	54
SANDSTRÖM, C. I.	1
SKEMP, R., with Kay, H.	153
SMITH, A. U., with Andjus, R. K., Knöpfelmacher, F., and Russell, R. W.	15
TOWNSEND, M. J., with Cheesman, G. H.	8
TROTTER, J. R.	97
TROTTER, J. R.	182
VAUTREY, P., with Fraisse, P.	114
WADE, E. A., with Littman, R. A.	186
WELFORD, A. T., with Hick, W. E.	39
YUDKIN, J., with Knöpfelmacher, F., Khairy, M., and Russell, R. W.	54
ZAJACZKOWSKA, A.	66
Corrigendum	143
ZANGWILL, O. L.	140

